Notes on Linear Algebra

by Oscar Bolina
E-mail: bolina@if.usp.br

This a set of homework assignments of a course in Linear Algebra I taught at the University of Califorina at Davis in the Spring of 2000.

The material presented here is taken from a number of sources, and I hope that references on page 146 will reflect this.

Many of the assignments bring the theory of the topics covered, while others contain a few problems only.

The following are included here:

1) Review problems on matrices and determinants.

2) The LU-decomposition.

3) Vectors, vector spaces and subspaces.

4) The subspaces associated with matrices.

5) The Gram-Schmidt orthogonalization and the QR-factorization.

6) The problem of least squares.

7) Eigenvalues and eigenvectors.

8) Linear transformations and similarity.

9) Schur's theorem.

10) Gerschgorin's theorem.

11) Computer exercises: Using Matlab.

12) Computer exercises: Systems of equations, rank, span and subspaces.

13) Computer exercises: Orthogonality.

14) Computer exercises: The least-squares problems, eigenvalues and the Gerschgorin's theorem.

15) A few extra problems.

1 Review problems on matrices and determinants.

1) In the matrix of letters $\mathbf{\Psi} = \begin{pmatrix} H & R & B & I \\ N & S & O & A \\ E & T & Y & L \\ M & G & D & I \end{pmatrix}$, spell out the words given by the following sequence of entries.

a) $a_{11}, a_{23}, a_{12}, a_{21}$

b) $a_{31}, a_{14}, a_{32}, a_{11}, a_{31}, a_{12}$

c) $a_{31}, a_{24}, a_{12}, a_{34}, a_{33}$

d) $a_{33}, a_{31}, a_{22}, a_{32}, a_{31}, a_{12}, a_{43}, a_{24}, a_{33}$

2) If $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, prove that $\mathbf{A}^2 - (a+d)\mathbf{A} + (ad - bc)\mathbf{I} = 0$.

3) Let $\mathbf{A} = \begin{pmatrix} 7 & 4 \\ -9 & -5 \end{pmatrix}$. Use induction to prove that, if $n \geq 1$,

$$\mathbf{A}^n = \begin{pmatrix} 1+6n & 4n \\ -9n & 1-6n \end{pmatrix}.$$

4) For the matrix $\mathbf{A} = \begin{pmatrix} 1/2 & \alpha \\ 0 & 1/2 \end{pmatrix}$, determine $\lim_{n\to\infty} \mathbf{A}^n$, where n is any integer. How does your answer depend on the value of α?

5) Let $\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 5 & 0 & 0 \end{pmatrix}$. Verify that $\mathbf{A}^3 = 5\mathbf{I}$ and find from this that $\mathbf{A}^{-1}$.

6) Show that if a square matrix $\mathbf{A}$ satisfies $\mathbf{A}^2 - 3\mathbf{A} + \mathbf{I} = 0$, then $\mathbf{A}^{-1} = 3\mathbf{I} - \mathbf{A}$.

7) Prove that if $\mathbf{A}$ is an invertible matrix, then its transpose $\mathbf{A}^T$ is also invertible and

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T.$$

8) Verify the following equalities, assuming that the stated inverses exist.

a) $(\mathbf{C}^{-1} + \mathbf{D}^{-1})^{-1} = \mathbf{C}(\mathbf{C} + \mathbf{D})^{-1}\mathbf{D}$

b) $(\mathbf{I} + \mathbf{CD})^{-1}\mathbf{C} = \mathbf{C}(\mathbf{I} + \mathbf{DC})^{-1}$

c) $(\mathbf{C} + \mathbf{DD}^T)^{-1}\mathbf{D} = \mathbf{C}^{-1}\mathbf{D}(\mathbf{I} + \mathbf{D}^T\mathbf{C}^{-1}\mathbf{D})^{-1}$

d) $(\mathbf{A} - \mathbf{B})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}(\mathbf{B}^{-1} - \mathbf{A}^{-1})\mathbf{A}^{-1}$

9) Verify that $\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 2 & 1 & 2 \end{pmatrix}$ satisfies $\mathbf{A}^3 = 3\mathbf{A}^2 - 3\mathbf{A} + \mathbf{I}$ and find $\mathbf{A}^{-1}$.

10) If $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices satisfying $\mathbf{A}^2 = \mathbf{B}^2 = (\mathbf{AB})^2 = \mathbf{I}$, prove that $\mathbf{AB} = \mathbf{BA}$.

11) Prove that if two symmetric matrices commute, then their product is symmetric.

12) Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices of the same order such that $\mathbf{A}$ commutes with $\mathbf{AB}$-$\mathbf{BA}$. Show that for any positive integer n,

$$\mathbf{A}^n\mathbf{B} - \mathbf{BA}^n = n(\mathbf{AB} - \mathbf{BA})\mathbf{A}^{n-1}.$$

13) Show that any two matrices that commute with $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ commute with each other.

14) A square matrix $\mathbf{A}$ is said to be *nilpotent* if $\mathbf{A}^p = 0$ for some positive integer p. If p is the least positive integer for which $\mathbf{A}^p = 0$, then $\mathbf{A}$ is said to be nilpotent of index p. Show that

$$\mathbf{A} = \begin{pmatrix} 1 & 5 & -2 \\ 1 & 2 & -1 \\ 3 & 6 & -3 \end{pmatrix}$$

is nilpotent of index 3.

15) A matrix is said to be involutory if $\mathbf{A}^2 = I$. Find all involutory matrices of size 2.

1.1 Partitioned Matrices.

Multiplication of partitioned matrices is accomplished by treating the submatrices derived from the partitioning as single elements. Consider the $m \times n$ $\mathbf{M}$ matrix that is partitioned into four submatrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$:

$$\mathbf{M}_{m\times n} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ -- & -- \\ \mathbf{C} & \mathbf{D} \end{array} \right).$$

Let $\mathbf{N}$ be a $p \times q$ matrix that is partitioned into four submatrices $\mathbf{E}$, $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$:

$$\mathbf{N}_{p\times q} = \left(\begin{array}{c|c} \mathbf{E} & \mathbf{F} \\ -- & -- \\ \mathbf{F} & \mathbf{H} \end{array} \right).$$

If $n = p$, then $\mathbf{M}$ and $\mathbf{N}$ are conformable for multiplication and the product $\mathbf{L} = \mathbf{M} \cdot \mathbf{N}$ is $m \times q$.
Now, if $\mathbf{A}$ is $j \times k$, then $\mathbf{B}$ is $j \times (n-k)$; if $\mathbf{C}$ is $(m-j) \times k$, then $\mathbf{D}$ is $(m-j) \times (n-k)$:

$$\mathbf{M}_{m\times n} = \left(\begin{array}{c|c} \mathbf{A}_{j\times k} & \mathbf{B}_{j\times(n-k)} \\ ---- & ---- \\ \mathbf{C}_{(m-j)\times k} & \mathbf{D}_{(m-j)\times(n-k)} \end{array} \right).$$

How should the submatrices $\mathbf{E}$, $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ be arranged in $\mathbf{N}$ so that the product

$$\mathbf{L} = \mathbf{MN} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ -- & -- \\ \mathbf{C} & \mathbf{D} \end{array} \right) \left(\begin{array}{c|c} \mathbf{E} & \mathbf{F} \\ -- & -- \\ \mathbf{G} & \mathbf{H} \end{array} \right)$$

can be computed treating the submatrices as if they were scalars? That is, when is it possible to write

$$\mathbf{L} = \mathbf{MN} = \left(\begin{array}{c|c} \mathbf{AE}+\mathbf{BG} & \mathbf{AF}+\mathbf{BH} \\ ---- & ---- \\ \mathbf{CE}+\mathbf{DG} & \mathbf{CF}+\mathbf{DH} \end{array} \right) \ ?$$

If in $\mathbf{N}$ the submatrix $\mathbf{E}$ is $r \times s$, then $\mathbf{F}$ is $r \times (q-s)$; if $\mathbf{G}$ is $(p-r) \times s$, then $\mathbf{H}$ is $(p-r) \times (q-s)$:

$$\mathbf{N}_{p\times q} = \begin{pmatrix} \mathbf{E}_{r\times s} & | & \mathbf{F}_{r\times(q-s)} \\ ---- & | & ---- \\ \mathbf{G}_{(p-r)\times s} & | & \mathbf{H}_{(p-r)\times(q-s)} \end{pmatrix}.$$

The product $\mathbf{AE}$ is well defined only if $k = r$,

$$\mathbf{A}_{j\times k} \cdot \mathbf{E}_{r\times s} = (\mathbf{AE})_{j\times s}.$$

Similarly, the product $\mathbf{BG}$ is well defined for $n - k = p - r$:

$$\mathbf{B}_{j\times(n-k)} \cdot \mathbf{G}_{(p-r)\times s} = (\mathbf{BG})_{j\times s}.$$

For this to be valid it is necessary that $n - k = p - r$. Since $k = r$, the condition is $n = p$. Thus there are only two requirements for the product $\mathbf{L} = \mathbf{MN}$ to be executed:

$$k = r \qquad \text{and} \qquad n = p.$$

We have then

$$\mathbf{L} = \mathbf{MN} = \begin{pmatrix} \mathbf{A}_{j\times k} & | & \mathbf{B}_{j\times(n-k)} \\ ---- & | & --- \\ \mathbf{C}_{(m-j)\times k} & | & \mathbf{D}_{(m-j)\times(n-k)} \end{pmatrix} \begin{pmatrix} \mathbf{E}_{k\times s} & | & \mathbf{F}_{k\times(q-s)} \\ ---- & | & ---- \\ \mathbf{G}_{(n-k)\times s} & | & \mathbf{H}_{(n-k)\times(q-s)} \end{pmatrix},$$

which means, as usual, that the number of columns of $\mathbf{A}$ equals the number of rows of $\mathbf{E}$, and similarly for the other matrix products.

16) If

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 2 & | & 3 & -1 \\ 0 & 1 & 1 & | & -1 & 2 \\ 3 & 2 & 1 & | & 2 & 3 \\ -- & -- & -- & | & -- & -- \\ -1 & 1 & 1 & | & 0 & 2 \end{pmatrix} = \begin{pmatrix} \mathbf{A} & | & \mathbf{B} \\ -- & | & -- \\ \mathbf{C} & | & \mathbf{D} \end{pmatrix}$$

and

$$\mathbf{N} = \begin{pmatrix} 4 & -1 & -1 & 0 & 0 \\ -1 & 3 & 0 & 0 & -1 \\ -1 & 0 & 4 & -1 & -1 \\ 0 & 0 & -1 & 3 & -1 \\ 0 & -1 & -1 & -1 & 4 \end{pmatrix} = \begin{pmatrix} \mathbf{E} & | & \mathbf{F} \\ -- & | & -- \\ \mathbf{G} & | & \mathbf{H} \end{pmatrix},$$

find an appropriate partitioned of $\mathbf{N}$ for the product $\mathbf{L} = \mathbf{MN}$ to be defined and find $\mathbf{L}$.

17) Compute $\mathbf{AB}$ using block multiplication, where

$$\mathbf{A} = \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & 0 \\ \hline 0 & 0 & 2 \end{array}\right) \quad \text{and} \quad \mathbf{B} = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 1 \\ \hline 0 & 0 & 0 & 1 \end{array}\right).$$

18) Use block multiplication to find $\mathbf{AB}$ where

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \end{pmatrix}.$$

(Hint: $\mathbf{A}$ and $\mathbf{B}$ can be written as partitioned matrices of the form $\mathbf{A} = \begin{pmatrix} \mathbf{C} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C} & \mathbf{C} \end{pmatrix}$.)

19) Use partitioned matrices to find $\mathbf{M}^2$, where

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -3 & 1 \end{pmatrix}$$

20) Find $\mathbf{A}^{300}$ for the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

(Hint: The matrix $\mathbf{C} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ is idempotent, that is, $\mathbf{C}^2 = \mathbf{C}$.)

21) Prove that if $\mathbf{M}$ is a matrix partitioned as $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ 0 & \mathbf{C} \end{pmatrix}$, and $\mathbf{A}$ and $\mathbf{C}$ are non-singular matrices, then

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{C}^{-1} \\ 0 & \mathbf{C}^{-1} \end{pmatrix}.$$

(If $\mathbf{B} = \mathbf{0}$ then $det\mathbf{M} = det\mathbf{A} \cdot det\mathbf{C}$. See problem 30.)

22) Find the inverse of the matrix $\begin{pmatrix} \mathbf{I} & 0 \\ \mathbf{A} & \mathbf{I} \end{pmatrix}$

1.2 Determinant facts

- If $\mathbf{A}$ is a square matrix of size n, the *minor* of entry a_{ij}, denoted by $M_{ij}(\mathbf{A})$, is the determinant of the submatrix of $\mathbf{A}$ obtained by deleting row i and column j of $\mathbf{A}$.

- The *cofactor* of entry a_{ij}, denoted by $\mathrm{cof}_{ij}(\mathbf{A})$, is given by $\mathrm{cof}_{ij}(\mathbf{A}) = (-1)^{i+j} M_{ij}(\mathbf{A})$.

- The determinant of an $n \times n$ matrix $\mathbf{A}$ can be computed by multiplying the entries in any row (or column) by their cofactors and adding the resulting products:

$$det(\mathbf{A}) = \sum_{i=1}^{n} a_{ij}\ \mathrm{cof}_{ij}(\mathbf{A}), \qquad j = 1, 2, ...n.$$

$$det(\mathbf{A}) = \sum_{j=1}^{n} a_{ij}\ \mathrm{cof}_{ij}(\mathbf{A}), \qquad i = 1, 2, ...n.$$

- If any two rows (or two columns) of a determinant are interchanged, then the sign of the value of the determinant is changed.

$$\begin{vmatrix} 2 & 6 & 4 \\ 1 & 3 & 0 \\ -1 & 0 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix} = 12$$

- If every element of a row of a determinant is zero, then the value of the determinant is zero.

- If every element of a row of a determinant is multiplied by the factor k, then the value of the determinant is multiplied by k.

$$\begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix} = 2\begin{vmatrix} 1 & 3 & 0 \\ 1 & 3 & 2 \\ -1 & 0 & 2 \end{vmatrix} = 6\begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ -1 & 0 & 2 \end{vmatrix} = 12\begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{vmatrix} = -12.$$

- The value of a determinant remains unchanged if every element of a row is increased by a scalar multiple of the corresponding element of another row.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21}+ka_{31} & a_{22}+ka_{32} & a_{23}+ka_{33} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

- If $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same size, then $det(\mathbf{BA}) = det(\mathbf{B}) \cdot \det(\mathbf{A})$.

23) Consider the matrix $\mathbf{A} = \begin{pmatrix} a & b & 0 & 1 \\ b & a & 0 & 0 \\ 0 & 0 & a & c \\ 1 & 0 & c & a \end{pmatrix}$ as an expression in a, b and c. Find the determinant of $\mathbf{A}$ by cofactor expansion.

24) Find the number k for which the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 & k \\ 3 & -1 & 1 \\ 5 & 3 & -5 \end{pmatrix}$ is singular.

25) Let $\mathbf{A}(x) = \begin{pmatrix} 2-x & 3 & 4 \\ 0 & 4-x & -5 \\ 1 & -1 & 3-x \end{pmatrix}$. Find $det(\mathbf{A})$, and then compute $d(det(\mathbf{A}))/dx$.

26) For the following tridiagonal matrix, $\mathbf{A}_n$, let $D_n = det(\mathbf{A})$, and derive the formula $D_n = 2D_{n-1} - D_{n-2}$ to deduce that $D_n = n+1$.

$$A_n = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}_{n\times n}.$$

27) Show that the **Vandermonde determinant** equals

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

(Hint: Take the following steps: Subtract the first column from the second, the third, ..., the n^{th} column; use expansion in cofactors along the first row of the resulting determinant; subtract from each row its preceding row multiplied by x_1; factor out the common factor $(x_i - x_1)$ from each column of the resulting determinant.)

28) Show that

$$\begin{vmatrix} 1 & a_1 & a_2 & \dots & a_n \\ 1 & a_1 + b_1 & a_2 & \dots & a_n \\ 1 & a_1 & a_2 + b_2 & \dots & a_n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_1 & a_2 & \dots & a_n + b_n \end{vmatrix} = b_1 b_2 \dots b_n.$$

29) Show that

$$\begin{vmatrix} a+b & b+c & c+a \\ p+q & q+r & r+p \\ u+v & v+w & w+u \end{vmatrix} = 2 \begin{vmatrix} a & b & c \\ p & q & r \\ u & v & w \end{vmatrix}$$

30) Find the determinant

$$D = \begin{vmatrix} a+d & 3a & b+2a & b+d \\ 2b & d+b & c-b & c-d \\ a+c & c-2d & d & a+3d \\ b-d & c-d & a+c & a+b \end{vmatrix}$$

by conveniently adding or subtracting columns 2, 3 and 4 to column 1.

31) If $\mathbf{A}_n$ is the tridiagonal matrix having 1's on the diagonal and sub- and super-diagonals,

$$\mathbf{A}_n = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 1 & 1 & 1 & \ldots & 0 & 0 & 0 \\ 0 & 1 & 1 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & 1 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 1 & 1 \end{pmatrix},$$

find $det(\mathbf{A}_n)$. (Hint: Use expansion in cofactors along the first column to obtain a relationship between $det(\mathbf{A}_n)$, $det(\mathbf{A}_{n-1})$ and $det(\mathbf{A}_{n-2})$.

32) Let $\mathbf{M}$ be a block matrix that is decomposed in the following way

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ 0 & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ 0 & \mathbf{I} \end{pmatrix},$$

where $\mathbf{A}$ and $\mathbf{D}$ are square matrices. Show that

$$det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = det(\mathbf{A}) \cdot det(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})$$

when $\mathbf{A}^{-1}$ exists.
We also have that

$$det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = det(\mathbf{D}) \cdot det(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})$$

when $\mathbf{D}^{-1}$ exists.

33) Show that if $a \neq b$ then

$$\begin{vmatrix} a+b & ab & 0 & \ldots & 0 & 0 \\ 1 & a+b & ab & \ldots & 0 & 0 \\ 0 & 1 & a+b & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a+b & ab \\ 0 & 0 & 0 & \ldots & 1 & a+b \end{vmatrix} = \frac{a^{n+1} - b^{n+1}}{a-b}.$$

What if $a = b$?

34) Find the $(n+2) \times (n+2)$ determinant

$$\begin{vmatrix} 0 & 1 & 1 & 1 & ... & 1 \\ 1 & 0 & a_1 & a_2 & ... & a_n \\ 1 & a_1 & 0 & a_1 + a_2 & ... & a_1 + a_n \\ 1 & a_2 & a_2 + a_1 & 0 & ... & a_2 + a_n \\ ... & ... & ... & ... & ... & ... \\ 1 & a_n & a_n + a_1 & a_n + a_2 & ... & 0 \end{vmatrix}.$$

2 The LU-decomposition.

2.1 Definition of triangular matrix.

A square matrix in which all entries below the main diagonal are zero is called **upper triangular**, and a square matrix in which all entries above the main diagonal are zero is called **lower triangular**.

1) a) Apply row operations to the matrix $A = \begin{pmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{pmatrix}$ to convert it to an upper triangular matrix.

b) Find the determinant of $\mathbf{A}$ from the upper triangular matrix obtained. (Remember to keep track of the row operations you perform!)

2.2 Definition of pivot.

A **pivot** of a nonzero row of a matrix is the first nonzero entry in that row. For example, the pivots of the matrix

$$\begin{pmatrix} 0 & 3 & 0 \\ 5 & 5 & 8 \\ 0 & 0 & 0 \end{pmatrix}$$

are the numbers 3 and 5. The bottom row, being entirely zero, has no pivot. We talk about pivots mainly in the conversion of a matrix to echelon form. For instance, the matrix in echelon form below

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -4 \\ 0 & 0 & -4 \end{pmatrix}$$

has pivots 1, -1 and -4.

2.3 The LU-decomposition of a matrix.

Let A be an $n \times n$ non-singular matrix. Suppose that $\mathbf{A}$ is converted to an upper triangular matrix $\mathbf{U}$ <u>without using row operations that interchange rows</u>. Supose also that $\mathbf{U}$ has no zero pivots.

Example 1. The matrix $A = \begin{pmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{pmatrix}$ can be converted to an upper triangular matrix $\mathbf{U}$ by the following sequence of elementary row operations

$$R_2 \to R_2 + (\mathbf{-2}) \times R_1$$

$$R_3 \to R_3 + (\mathbf{-3}) \times R_1$$

$$R_3 \to R_3 + (\mathbf{-4}) \times R_1.$$

The matrix $\mathbf{U}$ thus obtained is

$$\begin{pmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{pmatrix} = \mathbf{U}.$$

Let us denote by $m_{21} = -2$, $m_{31} = -3$ and $m_{32} = -4$ the various multipliers that were used in the example above to produce the zeros at positions $(2,1), (3,1)$ and $(3,2)$ of $\mathbf{U}$. Let us use these multipliers to construct a lower triangular matrix $\mathbf{L}$ that has 1's on its diagonal and the numbers $-m_{21}, -m_{31}$ and $-m_{32}$ at positions $(2,1), (3,1)$ and $(3,2)$:

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ -m_{31} & -m_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix}.$$

Then the matrix $\mathbf{A}$ can be written as the product

$$\mathbf{A} = \mathbf{LU}.$$

The above factorization of a square matrix $\mathbf{A}$ is called an **LU-decomposition** or **triangular decomposition** of $\mathbf{A}$.

2) Find the LU-decomposition of the matrix $\mathbf{A} = \begin{pmatrix} 3 & 2 & -1 \\ 9 & 7 & -5 \\ -6 & 6 & 3 \end{pmatrix}$.

3) Find the LU-decomposition of the matrix $\mathbf{A} = \begin{pmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{pmatrix}$.

Note that not all non-singular matrices have a LU-decomposition. In general, if row interchanges are required to convert $\mathbf{A}$ to upper triangular form, then there is no LU-decomposition of $\mathbf{A}$.

Another important question about LU-decomposition deals with the uniqueness of the decomposition. LU-decompositions are not unique. For example, it is equally possible to find a decomposition in which the diagonal elements of $\mathbf{U}$ are ones, instead of those of $\mathbf{L}$.

For example check that the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 5 & 8 \\ 4 & 9 & 10 \end{pmatrix}$ can be written as the product

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 5 & 8 \\ 4 & 9 & 10 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{pmatrix}.$$

by determining the constants l_{ij} and u_{ij}?

The important point about the LU-decomposition is that although there are many matrices $\mathbf{L}$ and $\mathbf{U}$ such that $\mathbf{A}$=$\mathbf{LU}$, both $\mathbf{L}$ and $\mathbf{U}$ are uniquely determined if we impose that $\mathbf{U}$ is upper-triangular with $u_{ii} \neq 0$ and $\mathbf{L}$ is lower-triangular with $l_{ii} = 1$.

4) Find the **LU**-decomposition for

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ -1 & 2 & -2 & 0 & 0 \\ 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 3 & -1 & 1 \\ 0 & 0 & 0 & -3 & 2 \end{pmatrix}.$$

2.4 Solving linear systems by factorization

Here are the steps to take in order to solve the linear system $\mathbf{Ax} = \mathbf{b}$ using the LU-decomposition. Assume that $\mathbf{A}$ is an $n \times n$ matrix, $\mathbf{x}$ and $\mathbf{b}$ are $n \times 1$ column vectors.

a) Find the LU-decomposition for $\mathbf{A}$ and rewrite the system as

$$\mathbf{Ax} = \mathbf{LUx} = \mathbf{b} \tag{1}$$

b) Introduce a new $n \times 1$ matrix $\mathbf{y}$ defined as

$$\mathbf{Ux} = \mathbf{y} \tag{2}$$

c) The system $\mathbf{LUx} = \mathbf{b}$ in (1) becomes

$$\mathbf{Ly} = \mathbf{b} \tag{3}$$

d) Solve the system (3) for $\mathbf{y}$, susbstitute it in (2) and then solve for $\mathbf{x}$.

5) Use the LU-decomposition to solve the system

$$\begin{aligned} x + 2y + 3z &= 16 \\ 3x + 5y + 8z &= 43 \\ 4x + 9y + 10z &= 57. \end{aligned}$$

6) Use the LU-decomposition of $\mathbf{A}$ given in *Example 1* to solve $\mathbf{Ax}$=$\mathbf{b}$, where

$$A = \begin{pmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 12 \\ 24 \\ 12 \end{pmatrix}.$$

7) Use the LU-decomposition to solve the system

$$\begin{aligned} x + 3y + 2z &= 19 \\ 2x + y + z &= 13 \\ 4x + 2y + 3z &= 31. \end{aligned}$$

3 Vectors, vector spaces and subspaces.

Definition of vector space. A nonempty set of vectors $\mathcal{V}$ is called a *vector space* (or a *linear space*) when the vector addition and scalar multiplication operations satisfy the following properties.

A1 The closure property for vector addition: $\mathbf{x}+\mathbf{y} \in \mathcal{V}$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$.

A2 $(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$ for each $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$.

A3 $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$ for each $\mathbf{x}, \mathbf{y} \in \mathcal{V}$.

A4 There is an element $\mathbf{0} \in \mathcal{V}$ such that $\mathbf{x}+\mathbf{0}=\mathbf{x}$ for each $\mathbf{x} \in \mathcal{V}$.

A5 For each $\mathbf{x} \in \mathcal{V}$, there is an element $-\mathbf{x} \in \mathcal{V}$ such that $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$.

M1 The closure property for scalar multiplication: $\alpha\mathbf{x} \in \mathcal{V}$ for each scalar α and $\mathbf{x} \in \mathcal{V}$.

M2 $(\alpha\beta)\mathbf{x}=\alpha(\beta\mathbf{x})$ for each scalar α and β, and $\mathbf{x} \in \mathcal{V}$.

M3 $\alpha(\mathbf{x}+\mathbf{y})=\alpha\mathbf{x}+\alpha\mathbf{y}$ for each scalar α, and $\mathbf{x}, \mathbf{y} \in \mathcal{V}$.

M4 $(\alpha+\beta)\mathbf{x}=\alpha\mathbf{x}+\beta\mathbf{x}$ for each scalar α and β, and $\mathbf{x} \in \mathcal{V}$.

M5 $\mathbf{1x}=\mathbf{x}$ for each $\mathbf{x} \in \mathcal{V}$.

This definition of vector addition and scalar multiplication stems from the addition and scalar multiplication that are familiar from our physical notion of vectors such as forces, velocities and accelerations. Thus, it is clear that the set of all vectors with n components is a vector space. However, many other algebraic systems permit definitions of addition and scalar multiplication that satisfy these same properties, even if the familiar notion of vector is not present at all. In this sense, any kind of object can be a vector, and the operations of addition and scalar multiplication may not have any relationship or similarity to the standard vector operations in R^n.
Here are some examples.

a) The set of polynomials in x with real coefficients is an example of a vector space with the above properties. A typical polynomial is

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0,$$

where the a_i are real numbers, and possibly zero. Two such polynomial can be added

$$\begin{aligned}(a_n x^n + \ldots + a_1 x + a_0) + (b_n x^n + \ldots &+ b_1 x + a_0) = \\ (a_n + b_n) x^n &+ \ldots + (a_1 + b_1) x + (a_0 + b_0)\end{aligned}$$

and multiplied by a scalar α

$$\alpha(a_n x^n + \ldots + a_1 x + a_0) = (\alpha a_n) x^n + \ldots + (\alpha a_1) x + (\alpha a_0)$$

b) Another example is the set of all solutions of the homogeneous system of equations $\mathbf{Ax} = \mathbf{0}$. This set is called the *solution space* or the *null space* of $\mathbf{A}$. This is one of the vector spaces associated with a matrix $\mathbf{A}$ which we will study in this course.

c) The set of all 2×2 matrices with real entries is a vector space if vector addition is defined to be matrix addition and scalar multiplication is defined to be matrix scalar multiplication.

d) The set of real-valued functions defined on the entire real line $(-\infty, \infty)$ is a vector space with respect to the operations of addition, $(f+g)(x) = f(x) + g(x)$, and multiplication by a scalar k, $(kf)(x) = kf(x)$.

e) The zero vector space $\mathcal{V} = \mathbf{0}$, where $\mathbf{0}$ is the zero vector is a vector space. Addition and scalar multiplication are given by the rules $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $k\mathbf{0} = \mathbf{0}$. The vector space axioms **A1**-**M5** are all true for $\mathcal{V}$.

Many times we are interested in examining a subset of a vector space that possesses the same structure as the vector space under consideration. This leads to the following definition.

3.1 Definition of Subspace.

A subset $\mathcal{W}$ of a vector space $\mathcal{V}$ is called a subspace of $\mathcal{V}$ if the following three conditions hold

i) $\mathbf{0} \in \mathcal{W}$

ii) $\mathbf{x} + \mathbf{y} \in \mathcal{W}$ whenever $\mathbf{x}, \mathbf{y} \in \mathcal{W}$.

iii) $\alpha\mathbf{x} \in \mathcal{W}$ for $\mathbf{x} \in \mathcal{W}$.

Here is an example of a subspace.

> Let n be a nonnegative integer, and let $p_n(x)$ consist of all polynomials in x with real coefficients having degree less than or equal to n. This is a subset of the set of polynomials defined in **a)** above. Note that the zero polynomial is an element of $p_n(x)$. Also, the sum of two polynomials with degree less than or equal to n is a polynomial of degree less than or equal to n, and the product of a scalar and polynomial of degree less than or equal to n is again a polynomial of degree less than or equal to n. So, the set $p_n(x)$ has the zero element, and is closed under addition and scalar multiplication. Therefore it is a subspace.

Thus, in order determine whether a subset is a subspace we just need to check whether the three conditions **i)**, **ii)** and **iii)** above hold. Here is an example of how to do it.

Example 1. Is the following subset

$$\mathcal{W} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix}, \quad xy \geq 0 \right\}$$

a subspace of R^2?

Solution. The set $\mathcal{W}$ will not be a subspace of R^2 if any one of conditions **i)**, **ii)** or **iii)** fail. So, all we need to do to prove that $\mathcal{W}$ is not a subspace is to find a counterexample to even one of these conditions. Now the vectors that comprise the subset $\mathcal{W}$ are in the

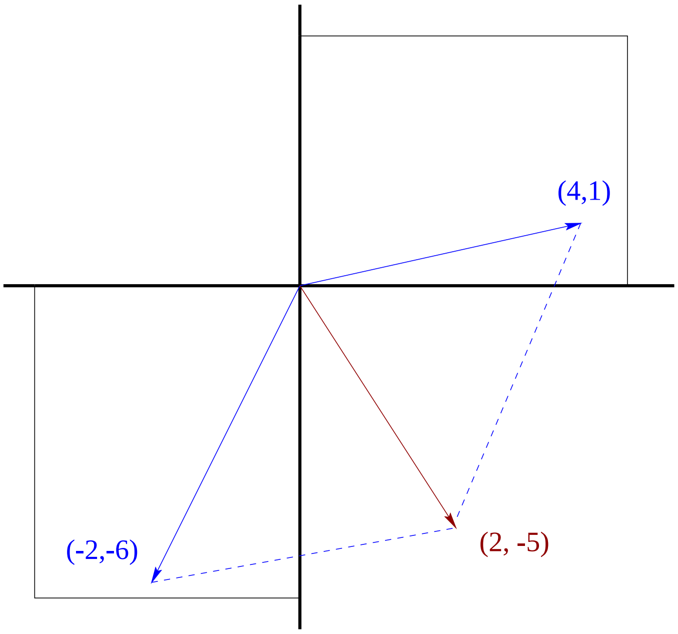

Figure 1: A subset of R^2 that is NOT a subspace of R^4

first and third quadrants of R^2, as shown in the figure below. Two elements of $\mathcal{W}$, namely,

$$\mathbf{v}_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} -2 \\ -6 \end{pmatrix}$$

are shown in Fig. 1. Observe that their sum $\mathbf{v}_1 + \mathbf{v}_2$ is NOT an element of $\mathcal{W}$ because it is a vector in the fourth quadrant, for which $xy < 0$. This is enough to prove that $\mathcal{W}$ is not a subspace.

8) Which of the following subsets of R^2 are subspaces?

a) $\mathcal{W}_1 = \{\begin{pmatrix} x \\ y \end{pmatrix}, \quad x = 2y.\}$

b) $\mathcal{W}_2 = \{\begin{pmatrix} x \\ y \end{pmatrix}, \quad x = 2y, \ 2x = y\}.$

c) $\mathcal{W}_3 = \{\begin{pmatrix} x \\ y \end{pmatrix}, \quad x = 2y + 1\}.$

d) $\mathcal{W}_4 = \{\begin{pmatrix} x \\ y \end{pmatrix}, \quad xy = 0\}.$

9) Let W be the subset of R^2 defined by

$$W = \{\mathbf{x} : \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \text{where either } x = 0 \text{ or } y = 0\}.$$

Is W a subspace of R^2.

10) Is the subset $\mathcal{W}$ of R^4 defined by

$$\mathcal{W} = \{\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}, \ 3x + y = 0, x + y + z = w\}$$

a subspace of R^4?

3.2 Definition of linear combination.

A vector $\mathbf{V}$ is said to be a linear combination of the vectors $\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p$ if and only if there exist scalars $c_1, c_2, ..., c_p$ such that

$$\mathbf{V} = c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + ... + c_p\mathbf{V}_p.$$

Example 2. Consider the vectors

$$\mathbf{V}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{V}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

in R^2. Show that the vector

$$\mathbf{b} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

can be written as a linear combination of $\mathbf{V}_1$ and $\mathbf{V}_2$.

Solution. If $\mathbf{b}$ is a linear combination of the given vectors, then

$$\mathbf{b} = c_1\mathbf{V}_1 + c_2\mathbf{V}_2 \tag{4}$$

for some constants c_1 and c_2. Explicitly (4) is

$$\begin{pmatrix} 4 \\ 3 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{5}$$

But note that (5) is the following matrix equation in disguise

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

Thus, the problem of determining a linear combination of vectors is equivalent to solving the system of linear equations $\mathbf{Ac} = \mathbf{b}$ for the $\mathbf{c}$, where

$$\mathbf{A} = [\mathbf{V}_1 \quad \mathbf{V}_2] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

is a matrix whose columns are the vectors $\mathbf{V}_1$ and $\mathbf{V}_2$. If the system is consistent, then there exist constants c_1 and c_2 such that (4) holds and $\mathbf{b}$ is a linear combination of given vectors. If the system is inconsistent, and thus admits no solution, then $\mathbf{b}$ can not be written in terms of the two given vectors. In this example the system is consistent and the unique solution can easily be seen to be $c_2 = 3$ and $c_1 = 1$.

3.3 Definition of linear dependency.

Let $\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p$ be p n-dimensional vectors. If there exist scalars $c_1, c_2, ..., c_p$, not all zero, such that

$$c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + ... + c_p\mathbf{V}_p = 0,$$

then the vectors $\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p$ are said to be linearly dependent; otherwise these vectors are said to be linearly independent.

Example 3. Determine whether the vectors

$$\mathbf{V}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{V}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \mathbf{V}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

are linearly dependent or linear independent.

Solution. The given vectors will be linearly dependent if

$$c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + c_3\mathbf{V}_3 = 0, \tag{6}$$

for some c_1, c_2, c_3, not all zero. But (6) is equivalent to the following system of equation for the constants c_1, c_2 and c_3:

$$c_1\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + c_2\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + c_3\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \mathbf{0}.$$

We have thus converted the problem of determining linear dependency among vectors into one of solving a (homogeneous) system of linear equations $\mathbf{Ac} = \mathbf{0}$ for the constant c_1, c_2 and c_3, where the coefficient matrix $\mathbf{A}$ has as columns the vectors $\mathbf{V}_1, \mathbf{V}_2$ and $\mathbf{V}_3$. If the only solution is $c_1 = c_2 = c_3 = 0$, then the vectors $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$ are linearly independent; if the system admits non-trivial solutions, then these vectors are linearly dependent.
Reducing the coefficient matrix to echelon form

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ -1 & -1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

we obtain as solutions $c_1 = -c_3$ and $c_2 = c_3$ with c_3 is free. Thus, all solutions can be put in the form

$$\mathbf{c} = c_3\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

and the vectors $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$ are linearly dependent. You might want to check that the dependency relationship is

$$(\mathbf{-1}) \cdot \mathbf{V}_1 + (\mathbf{1}) \cdot \mathbf{V}_2 + (\mathbf{1}) \cdot \mathbf{V}_3 = \mathbf{0},$$

where the coefficients $-1, 1, 1$ are the entries of $\mathbf{c}$.

Example 4. Determine whether the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 5 \\ 6 \\ 4 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 2 \\ 6 \\ 8 \\ 5 \end{pmatrix}$$

are linearly dependent or independent in R^4.

Solution 1. We determine whether the vectors are linearly independent by writing down the coefficient matrix of the system $\mathbf{Ac} = \mathbf{0}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 2 & 5 & 6 \\ 1 & 3 & 6 & 8 \\ 1 & 2 & 4 & 5 \end{pmatrix}$$

has as columns the given vectors, and converting it to echelon form. We obtain

$$\mathbf{A} \longrightarrow \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This system has non-trivial solutions. The set is linearly dependent.

Solution 2. There is a second solution to this problem that reveals important facts about matrices and their reduced row-echelon form. Suppose that we want to determine the dependency relationship – if one exists – between the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and $\mathbf{v}_4$. In other words, suppose we want to express some of the given vectors as a linear combination of the others. To do this, convert the matrix $\mathbf{A}$ to reduced row-echelon form. We have

$$\mathbf{A} \longrightarrow \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 & \mathbf{w}_4 \end{pmatrix},$$

where we have denoted the **column** vectors by $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ and $\mathbf{w}_4$. Note that from this arrangement it is easy to see that $\mathbf{w}_1, \mathbf{w}_2$ and $\mathbf{w}_3$ are linearly independent vectors and that $\mathbf{w}_4$ can be written as the linear combination

$$\mathbf{w}_4 = -\mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3. \tag{7}$$

The question is: What does this have to do with the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and $\mathbf{v}_4$ of our original problem? The answer is that the same relationship (7) that exists between the $\mathbf{w}$ vectors holds for the $\mathbf{v}$ vectors. Thus we can confidently write (Check it yourself!)

$$\mathbf{v}_4 = -\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3.$$

11) Determine whether the vectors

$$X_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix} \qquad X_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix} \qquad X_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix}$$

are linearly independent vectors in R^4.

12) Find the value of k such that the vectors

$$\mathbf{V}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{V}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{V}_3 = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{V}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ k \end{pmatrix}$$

are linearly independent.

13) For which real numbers λ are the following vectors linearly independent?

$$X_1 = \begin{pmatrix} \lambda \\ -1 \\ -1 \end{pmatrix} \qquad X_2 = \begin{pmatrix} -1 \\ \lambda \\ -1 \end{pmatrix} \qquad X_3 = \begin{pmatrix} -1 \\ -1 \\ \lambda \end{pmatrix}.$$

14) Let $\mathbf{a}$ and $\mathbf{b}$ be linearly <u>independent</u> vectors in R^2. Find the value of α making the vectors $\alpha\mathbf{a} + 2\mathbf{b}$ and $\mathbf{a} - \mathbf{b}$ linearly <u>dependent</u>.

15) Let **a**, **b** and **c** be three linearly <u>independent</u> vectors in R^3. For what values of α are the vectors $\alpha\mathbf{a}+4\mathbf{b}+2\mathbf{c}$ and $\mathbf{a}+\alpha\mathbf{b}-\mathbf{c}$ linearly <u>dependent</u>?

3.4 Definition of span.

If $\{\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p\}$ is a set of vectors in a vector space $\mathcal{V}$, then the subspace $\mathcal{W}$ of $\mathcal{V}$ consisting of all linear combinations

$$c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + ... + c_p\mathbf{V}_p$$

of the vectors $\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p$ is called the space spanned by $\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p$. We also say that $\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p$ span $\mathcal{W}$, and represent it by writing

$$\mathcal{W} = span\{\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p\}.$$

Example 5. The span of the vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

is the plane R^2, since every vector in that plane can be written as a linear combination

$$\mathbf{v} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2.$$

We say the set $\{\mathbf{e}_1, \mathbf{e}_2\}$ spans R^2.

Example 6. Let $\mathcal{W}$ be the subspace of R^4 spanned by the vectors

$$\mathbf{V}_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 3 \end{pmatrix} \qquad \mathbf{V}_2 = \begin{pmatrix} 2 \\ -5 \\ -3 \\ 6 \end{pmatrix} \qquad \mathbf{V}_3 = \begin{pmatrix} 2 \\ -1 \\ 4 \\ -7 \end{pmatrix}.$$

Determine whether the vector

$$\mathbf{b} = \begin{pmatrix} 5 \\ -8 \\ 1 \\ 2 \end{pmatrix}.$$

belongs to $\mathcal{W}$.

Solution. The vector $\mathbf{b}$ belongs to $\mathcal{W}$ if it can be written as a linear combination of the vectors in $\mathcal{W}$,

$$\mathbf{b} = c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + c_3\mathbf{V}_3, \tag{8}$$

for some constants c_1, c_2 and c_3. But solving (8) for the c's is equivalent to solving the system $\mathbf{Ac} = \mathbf{b}$, where the coefficient matrix $\mathbf{A}$ has as columns the vectors $\mathbf{V}_1, \mathbf{V}_2$ and $\mathbf{V}_3$:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 \\ -2 & -5 & -1 \\ 0 & -3 & 4 \\ 3 & 6 & -7 \end{pmatrix}.$$

Thus, (8) becomes the matrix equation

$$\begin{pmatrix} 1 & 2 & 2 \\ -2 & -5 & -1 \\ 0 & -3 & 4 \\ 3 & 6 & -7 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -8 \\ 1 \\ 2 \end{pmatrix}. \tag{9}$$

Now, the vector $\mathbf{b}$ lies in the subspace $\mathcal{W}$ if and only if $\mathbf{Ac} = \mathbf{b}$ has a solution. To find a solution to the system above, convert its augmented matrix to echelon form. We obtain

$$\left(\begin{array}{ccc|c} 1 & 2 & 2 & 5 \\ -2 & -5 & -1 & -8 \\ 0 & -3 & 4 & 1 \\ 3 & 6 & -7 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 2 & 5 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right). \tag{10}$$

The system is consistent and has a unique solution $c_1 = c_2 = c_3 = 1$. Therefore $\mathbf{b} = \mathbf{V}_1 + \mathbf{V}_2 + \mathbf{V}_3$ and $\mathbf{b}$ is in $\mathcal{W}$.

In the above example we have tested whether an arbitrary vector belonged to a subspace. Some vectors will belong to a particular subspace and others will not. In *example 8* below we determine <u>all</u> the vectors $\mathbf{b}$ that belong to a subspace.

Example 8. Give a characterization of of the subspace $\mathcal{W}$ spanned by the vectors

$$\mathbf{V}_1 = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix} \qquad \mathbf{V}_2 = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}.$$

Solution. The criterium for membership in $\mathcal{W}$ is equation (8), which translates itself into determining whether the equation (9) is consistent for an arbitrary vector $\mathbf{b}$. Equation (10) becomes

$$\begin{pmatrix} 1 & 2 & | & b_1 \\ -3 & -1 & | & b_2 \\ 2 & -1 & | & b_3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & | & b_1 \\ 0 & 5 & | & b_2 + 3b_1 \\ 0 & 0 & | & b_1 + b_2 + b_3 \end{pmatrix}.$$

This system is consistent if and only if $b_1 + b_2 + b_3 = 0$. Thus, $\mathcal{W}$ is the subspace

$$\mathcal{W} = \left\{ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \text{ such that } b_1 + b_2 + b_3 = 0 \right\}.$$

In the problems below you have to determined whether the given vectors span R^3. This means that you have to determine whether an arbitrary vecotr $\mathbf{b}$ in R^3 can be written as a linear combination of the given vectors. To do this, place the given vectors as column in a matrix $\mathbf{A}$ and determine whether the system $\mathbf{Ax} = \mathbf{b}$ is consistent for every $\mathbf{b} \in R^3$. If your answer is affirmative, the given set spans R^3.

16) Verify whether or not

$$\mathcal{S} = \{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \}$$

spans R^3.

17) Verify whether or not

$$\mathcal{S} = \{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \}$$

spans R^3.

18) Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be the subspace of R^3 defined by

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 0 \\ -7 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 7 \\ 0 \end{pmatrix}.$$

Does S span R^3?

3.5 Definition of basis.

A *basis* of a vector space $\mathcal{V}$ is a linearly independent set of vectors $\{\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p\}$ which spans $\mathcal{V}$. The dimension of a vector space $\mathcal{V}$ is defined to be the number of vectors in any basis for $\mathcal{V}$.

Example 9. The real vector space R^3 has basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

To prove this, we need to show that 1) the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are linearly independent and 2) any vector

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

in R^3 can be written as a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. We leave the (easy) proof of the first part as an exercise. The proof of the second part follows from the fact that

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3.$$

Example 10. Determine whether the set of vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 5 \\ 6 \\ 4 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 2 \\ 6 \\ 8 \\ 5 \end{pmatrix}$$

forms a basis of R^4.

Solution. The given vectors form a basis if they are linearly independent. But we have already determined in *Example 4* that the set is linearly dependent and therefore does not form a basis of R^4.

19) Find a basis for the subspace of R^4 spanned by the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 2 \\ 3 \\ 2 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 3 \\ 3 \\ 3 \\ 3 \end{pmatrix}, \quad \mathbf{v}_5 = \begin{pmatrix} 5 \\ 3 \\ 5 \\ 3 \end{pmatrix}.$$

20) Find a basis for the subspace of R^4 spanned by the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 1 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

21) Let W be a subset of R^4 spanned by the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ -1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 4 \\ -1 \\ 5 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 4 \\ -1 \end{pmatrix}, \quad \mathbf{v}_5 = \begin{pmatrix} 2 \\ 5 \\ 0 \\ 2 \end{pmatrix}.$$

Find a subset that is a basis for W.

22) Find the dimension of the subspace of R^5 that is spanned by the following vectors

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 5 \\ -2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 3 \\ -2 \\ 5 \\ -1 \\ 3 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 5 \\ -3 \\ 8 \\ 4 \\ 1 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 11 \\ -7 \end{pmatrix}.$$

23) Answer true (T) or false (F).

a) If the vector $\mathbf{0}$ is included among the vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ then these vectors are linearly dependent.

b) If $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ are linearly independent and $\mathbf{v}_{p+1}$ is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$, then the vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_{p+1}$ are linearly independent.

c) If any $p-1$ vectors of $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ are linearly independent, then $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ are linearly independent.

d) If $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ are linearly independent then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_p\mathbf{v}_p = 0$ implies $c_1 = c_2 = ... = c_p = 0$.

e) If $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ are linearly independent then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_p\mathbf{v}_p = 0$ is possible **only** if $c_1 = c_2 = ... = c_p = 0$.

f) If $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ are linearly independent then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_p\mathbf{v}_p = 0$ is never true except when $c_1 = c_2 = ... = c_p = 0$.

g) If $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ are linearly independent then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_p\mathbf{v}_p = 0$ has **no** non-trivial solutions.

h) $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p$ are linearly independent means that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_p\mathbf{v}_p = 0$ has the trivial solution.

i) The equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_p\mathbf{v}_p = 0$ is linearly independent.

4 The subspaces associated with matrices.

In this assignment we are going to use the ideas developed in the previous assignment to study four fundamental subspaces associated with a matrix $\mathbf{A}$. For an $m \times n$ matrix $\mathbf{A}$ these are

a. The column space of $\mathbf{A}$, $col(\mathbf{A})$, is the space spanned by the columns of $\mathbf{A}$.

b. The row space of $\mathbf{A}$, $row(\mathbf{A})$, is the space spanned by the rows of $\mathbf{A}$.

c. The null space of $\mathbf{A}$, $null(\mathbf{A})$, is the set of all solutions to the homogeneous system $\mathbf{Ax} = \mathbf{0}$.

d. The null space of $\mathbf{A}^T$ is the set of all solutions to the homogeneous system $\mathbf{A}^T\mathbf{x} = \mathbf{0}$.

4.1 The column space.

For an $m \times n$ matrix $\mathbf{A}$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

the vectors

$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ . \\ a_{m1} \end{pmatrix}, \quad \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ . \\ a_{m2} \end{pmatrix}, \quad \dots, \quad \mathbf{c}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ . \\ a_{mn} \end{pmatrix}$$

formed from the columns of $\mathbf{A}$ are called the column vectors of $\mathbf{A}$. The subspace of R^m spanned by the column vectors of $\mathbf{A}$ is called the **column space** of $\mathbf{A}$, usually denoted by $col(\mathbf{A})$. The dimension of this subspace is the number of linearly independent column vectors of $\mathbf{A}$.

Example 1. Find a basis for the column space of

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 5 & 1 & 2 \\ 0 & 4 & -4 & 4 \end{pmatrix}.$$

Solution. The column vectors of $\mathbf{A}$ are the columns of $\mathbf{A}$ viewed as vectors. These vectors may be linearly dependent or independent. To find a basis for the column space of $\mathbf{A}$, convert $\mathbf{A}$ to reduced row-echelon form. We have

$$\mathbf{A} = \begin{array}{c} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 5 & 1 & 2 \\ 0 & 4 & -4 & 4 \end{pmatrix} \\ \begin{array}{cccc} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \end{array} \end{array} \to \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{array}{c} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ \begin{array}{cccc} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 & \mathbf{w}_4 \end{array} \end{array}.$$

From the reduced row-echelon form we can easily see that

$$\mathbf{w}_3 = 2\mathbf{w}_1 - \mathbf{w}_2 \quad \text{and} \quad \mathbf{w}_4 = -\mathbf{w}_1 + \mathbf{w}_2.$$

But we already know that this same relationship that holds for the vectors in the reduced row-echelon matrix also holds for the vectors $\mathbf{v}$ of $\mathbf{A}$. Thus we can automatically write down that

$$\mathbf{v}_3 = 2\mathbf{v}_1 - \mathbf{v}_2 \quad \text{and} \quad \mathbf{v}_4 = -\mathbf{v}_1 + \mathbf{v}_2.$$

This last equation clearly shows that $\mathbf{v}_1$ and $\mathbf{v}_2$ forms a basis for the column space, since the other column vectors can be written as a linear combination of these two. Because there are two column vectors that span the column space, the dimension of the column space is two.
Note the the point of transforming the given matrix to reduced row-echelon form is simply that in that form we can determine more easily the relationship between the column vectors. However, if we are not interested in the exact relationship between the columns but just want to find a basis for the subspace, then the reduced row-echelon (and also the row-echelon) form tells us that the linearly independent columns of $\mathbf{A}$ are the corresponding columns of the reduced row-echelon form that contain the leading ones.

1) Find a basis for the column space of

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{pmatrix}.$$

4.2 The row space.

For an $m \times n$ matrix $\mathbf{A}$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

the vectors

$$\begin{aligned} \mathbf{r}_1 &= (a_{11}, a_{12}, ..., a_{1n}) \\ \mathbf{r}_2 &= (a_{21}, a_{22}, ..., a_{2n}) \\ . \quad . &\quad . \quad . \quad . \quad . \quad . \quad . \\ \mathbf{r}_m &= (a_{m1}, a_{m2}, ..., a_{mn}) \end{aligned}$$

formed from the rows of $\mathbf{A}$ are called the row vectors of $\mathbf{A}$. The subspace of R^n spanned by the row vectors of $\mathbf{A}$ is called the **row space** of $\mathbf{A}$, usually denoted by $row(\mathbf{A})$.

Example 2. Find a basis for the row space of

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 5 & 1 & 2 \\ 0 & 4 & -4 & 4 \end{pmatrix}.$$

Solution 1. The row vectors of $\mathbf{A}$ are the rows of $\mathbf{A}$ viewed as vectors. These vectors may be linearly dependent or independent. To find a basis, convert the matrix $\mathbf{A}$ to row-echelon form $\mathbf{B}$. The nonzero rows of **B** form a basis for the row space of **A**. The dimension of this subspace is the rank of **A**. We have

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 5 & 1 & 2 \\ 0 & 4 & -4 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The vectors

$$\mathbf{r}_1 = (1,1,1,0) \quad \text{and} \quad \mathbf{r}_2 = (0,1,-1,1)$$

form a basis for the row space of **A**.

Solution 2. Convert the matrix **A** to reduced row-echelon form **B**. The nonzero rows of **B** form a basis for the row space of **A**. We have

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 5 & 1 & 2 \\ 0 & 4 & -4 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The vectors

$$\mathbf{r}_1 = (1,0,2,-1) \quad \text{and} \quad \mathbf{r}_2 = (0,1,-1,1)$$

form a basis for the row space of **A**.

Although the basis vectors just obtained differ from the basis vectors in the previous example, the two bases do span the same subspace (How would you check this?). This is so because of the following result

> If a matrix **B** is obtained from a matrix **A** by an elementary row operation, then the row space of **A** is the same as the row space of **B**.

This result shows that we can find bases for the row space of **A** by converting **A** to any of its row-echelon forms. They will span the same subspace. However, since the reduced row-echelon form of any matrix is unique, the basis formed from its non-zero rows is said to be the *canonical basis* of $row(\mathbf{A})$.
Observe that in the first solution the basis for the row space of **A**

contained just one row vectors of $\mathbf{A}$, while the second basis contained no row vector of $\mathbf{A}$. Another basis, consisting entirely of row vector of $\mathbf{A}$, can be obtained from the rows of $\mathbf{A}$ that correspond to nonzero rows of the echelon form $\mathbf{B}$. We have

$$\mathbf{r}_1 = (1,1,1,0) \quad \text{and} \quad \mathbf{r}_2 = (3,5,1,2).$$

2) Find a basis for the row space of

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{pmatrix}.$$

4.3 Using the transpose.

An elementary observation suggests a second way to find bases for row spaces and column spaces. Since the operation of transposing a matrix converts row vectors into column vectors and column vectors into row vectors, the row space of $\mathbf{A}^T$ is the same as the column space of $\mathbf{A}$, and the column space of $\mathbf{A}^T$ is the same as the row space of $\mathbf{A}$. Let us use this fact to redo the examples 1 and 2 above.

Example 1, Solution 2. We want to find a basis for the column space of

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 5 & 1 & 2 \\ 0 & 4 & -4 & 4 \end{pmatrix}.$$

The column space of $\mathbf{A}$ is the row space of $\mathbf{A}^T$. To find a basis for the column space of $\mathbf{A}$, **transpose** the matrix $\mathbf{A}$ and convert $\mathbf{A}^T$ to echelon form $\mathbf{B}$. The nonzero rows of $\mathbf{B}$ form a basis for row space of $\mathbf{A}^T$. They also form a basis for the column space of $\mathbf{A}$.

$$\mathbf{A}^T = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 5 & 4 \\ 1 & 1 & -4 \\ 0 & 2 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The row vectors of $\mathbf{A}^T$

$$\mathbf{r}_1 = (1,3,0) \qquad \text{and} \qquad \mathbf{r}_2 = (0,1,2)$$

form a basis for the column space of $\mathbf{A}$.

Example 2, Solution 3. We want to find a basis for the row space of

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 5 & 1 & 2 \\ 0 & 4 & -4 & 4 \end{pmatrix}.$$

We will transpose $\mathbf{A}$, thereby converting the row space of $\mathbf{A}$ into the column space of $\mathbf{A}^T$, and then we will transform $\mathbf{A}^T$ to reduced row-echelon form.

$$\mathbf{A}^T = \begin{array}{c} \begin{pmatrix} 1 & 3 & 0 \\ 1 & 5 & 4 \\ 1 & 1 & -4 \\ 0 & 2 & 4 \end{pmatrix} \\ \begin{array}{ccc} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{array} \end{array} \longrightarrow \begin{pmatrix} 1 & 3 & 0 \\ 0 & 2 & 4 \\ 0 & -2 & -4 \\ 0 & 2 & 4 \end{pmatrix} \longrightarrow \begin{array}{c} \begin{pmatrix} 1 & 0 & -6 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \begin{array}{ccc} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{array} \end{array}$$

From the reduced row-echelon form we see that the column vectors of $\mathbf{A}^T$ satisfy the relationship

$$\mathbf{v}_3 = -6\mathbf{v}_1 + 2\mathbf{v}_2$$

The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ form a basis for the column space of $\mathbf{A}^T$ and therefore also for the row space of $\mathbf{A}$. Note that in previous examples we determined bases for row spaces without finding an explicit relationship between the row vectors. The procedure outlined above has the advantage that it allows us to do just that.

3) Given the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{pmatrix},$$

use the transpose to find a relationship between the row vectors of $\mathbf{A}$.

4.4 The null space.

In Section 3, page 17, we said that the set of all solutions to the homogeneous system

$$\mathbf{Ax} = \mathbf{0},$$

where $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{x} \in R^n$, was one of four subspace associated with matrices. This subspace is called the **null** space of $\mathbf{A}$, usually denoted by

$$null(\mathbf{A}) = \{\mathbf{x} \in R^n \mid \mathbf{Ax} = \mathbf{0}\}.$$

The dimension of the null space of $\mathbf{A}$ is the number linearly independent vectors that comprise the solution of $\mathbf{Ax} = \mathbf{0}$.

Example 3. Find a basis for the null spaces of

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$$

Solution. When the augmented matrix of the system $\mathbf{Ax} + \mathbf{0}$ is converted to echelon form we obtain

$$\begin{pmatrix} 1 & 2 & | & 0 \\ 3 & 4 & | & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & | & 0 \\ 0 & -1 & | & 0 \end{pmatrix}.$$

This system has only the trivial solution $x = y = 0$. Thus, $null(\mathbf{A}) = \{\mathbf{x} : \mathbf{x} = \mathbf{0}\}$.

When the augmented matrix of the system $\mathbf{Bx} = \mathbf{0}$ is converted to echelon form we obtain

$$\begin{pmatrix} 1 & 3 & | & 0 \\ 2 & 6 & | & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & | & 0 \\ 0 & 0 & | & 0 \end{pmatrix}.$$

The solutions to this system are of the form

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} = y \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Thus,

$$null(\mathbf{B}) = \{\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix},\ x + 3y = 0\}.$$

and the vector $\mathbf{v} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$ is a basis for the null space of $\mathbf{B}$.

Example 4. Find the dimension of the null space of

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 0 & 4 \\ 3 & 1 & 1 & 0 \\ -1 & -5 & -1 & 8 \end{pmatrix}.$$

Solution. By converting the augmented matrix of the system to row-echelon form

$$\left(\begin{array}{cccc|c} 1 & -2 & 0 & 4 & 0 \\ 3 & 1 & 1 & 0 & 0 \\ -1 & -5 & -1 & 8 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & -2 & 0 & 4 & 0 \\ 0 & 7 & 1 & -12 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

we obtain as solution of the homogeneous system $\mathbf{Ax} = \mathbf{0}$ all vectors of the form

$$\mathbf{x} = x_3 \begin{pmatrix} -2 \\ -1 \\ 7 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -4 \\ 12 \\ 0 \\ 7 \end{pmatrix}.$$

The null space of $\mathbf{A}$ is the following set

$$null(\mathbf{A}) = \{ \begin{pmatrix} -2 \\ -1 \\ 7 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 12 \\ 0 \\ 7 \end{pmatrix} .\}$$

These two vectors form a basis for the null($\mathbf{A}$) and thus the dimension of the null space is 2.

4) Find a basis for the null space of

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 2 \\ 2 & 2 & 0 & 4 \\ 0 & 1 & 1 & 0 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

5) Find a basis for the null space of

$$\begin{pmatrix} 1 & -2 & 1 & -5 \\ 2 & 1 & 7 & 5 \\ 1 & -1 & 2 & -2 \end{pmatrix}.$$

6) Find a basis for the null space of $\begin{pmatrix} 1 & 0 & 5 & -2 & -1 \\ 0 & 1 & -2 & 4 & 2 \end{pmatrix}$.

4.5 The null space of $\mathbf{A}^T$.

Everything that we said about the null space of $\mathbf{A}$ applies for the null space of $\mathbf{A}^T$ with obvious changes. Since no new idea is involved, there is need to repeat the discussion here.

4.6 Rank and nullity.

The dimension of the null space of $\mathbf{A}$ is called the **nullity** of $\mathbf{A}$ and is denoted nullity($\mathbf{A}$). The dimension of the null space of $\mathbf{A}$ is related to the rank of $\mathbf{A}$. When we studied linear systems we defined the rank of a matrix as the number of nonzero rows of its echelon form or reduced row-echelon form. In *Example 2* we saw that the number of nonzero rows in the echelon form of a matrix is the dimension of its row space. Thus, the rank of $\mathbf{A}$ is also a measure of the dimension of a subspace, in this case of the row space of $\mathbf{A}$. In this section we are going to state and prove a theorem that relates the dimensions of the row and null spaces of $\mathbf{A}$.
In *Example 4* above, we saw that $nullity(\mathbf{A}) = 2$, and $rank(\mathbf{A}) = 2$, and therefore $nullity(\mathbf{A}) + rank(\mathbf{A}) = 4$, the number of columns of the matrix $\mathbf{A}$. This is an example of the following general result

> **Theorem.** If $\mathbf{A}$ is a matrix with n columns, then
> $$rank(\mathbf{A}) + nullity(\mathbf{A}) = n.$$

This theorem is called the *rank plus nullity theorem* or the *dimension theorem for matrices.*

Proof. Consider the system of equations $\mathbf{A}\mathbf{x} = \mathbf{0}$, where $\mathbf{A}$ is an $m \times n$ matrix of rank r. Suppose that the augmented matrix of the system has been reduced to echelon form. Since the rank of the augmented matrix is also r, there are $n - r$ free variables in the solution $\mathbf{x}$ of $\mathbf{A}\mathbf{x} = \mathbf{0}$. This is exactly the number of dimensions of the null space of $\mathbf{A}$. Thus, $nullity(\mathbf{A}) = n - r$. Since $r = rank(\mathbf{A})$, it follows that $rank(\mathbf{A}) + nullity(\mathbf{A}) = n$.

7) Find a basis for the row, column and null spaces of the following matrices

a) $\begin{pmatrix} 1 & 1 & 5 & 1 & 4 \\ 2 & -1 & 1 & 2 & 2 \\ 3 & 0 & 6 & 0 & -3 \end{pmatrix}$ **b)** $\begin{pmatrix} 1 & 1 & 2 & 0 & 1 \\ 2 & 2 & 5 & 0 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 8 & 11 & 19 & 0 & 11 \end{pmatrix}$

and verify the rank plus nullity theorem.

8) Use the algorithm (arrow diagram **3.5.2** of the textbook [4]) to find the bases for the null, column and row spaces of

a) $\begin{pmatrix} 1 & -1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 \\ 3 & -2 & 4 & 1 & 1 \end{pmatrix}$ **b)** $\begin{pmatrix} 2 & 4 & 1 \\ 0 & -2 & -1 \\ 1 & 3 & 1 \end{pmatrix}$.

9) Use the algorithm (arrow diagram **3.6.4** of the textbook [4]) to determine whether the following systems of equations are consistent and solve them in case they are consistent.

a) $\begin{pmatrix} 1 & 2 & -2 & 1 & -1 \\ 2 & 3 & 1 & -1 & 2 \\ 3 & 5 & -1 & 0 & 1 \\ 1 & 1 & 3 & -2 & 3 \end{pmatrix} \mathbf{x} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$.

b) $\begin{pmatrix} 1 & 2 & -1 \\ 1 & 3 & 1 \\ 3 & 8 & 4 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ 5 \\ 17 \end{pmatrix}$.

10) Answer true (T) or false (F).

a) A basis for a vector space $\mathcal{V}$ is a set of vectors that is linearly independent and also spans $\mathcal{V}$.

b) A vector space is finite dimensional if it is a basis containing a finite number of vectors.

c) If $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p\}$ is a basis for a vector space $\mathcal{V}$, then every linearly independent set of vectors in $\mathcal{V}$ must contain p or fewer vectors.

d) Every basis for a finite-dimensional vector space must contain the same number of vectors.

e) The dimension of a vector space is the number of vectors in a basis for that vector space.

f) The dimension of the row space of a matrix $\mathbf{A}$ is equal to the rank of $\mathbf{A}$.

g) If $\mathbf{A}$ is any matrix, then the row space and the column space of $\mathbf{A}$ have the same dimension.

h) If $\mathbf{A}$ is a 5×7 matrix with rank 4, and if $\mathbf{Ax} = \mathbf{b}$ is a consistent linear system, then the general solution of the system contains $7 - 4 = 3$ free variables.

i) Elementary row operations do not change the row space of a matrix.

j) Elementary row operations do not change the null space of a matrix.

k) Elementary row operations do change the column space of a matrix.

l) A system of linear equations $\mathbf{Ax} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of $\mathbf{A}$.

5 The Gram-Schmidt orthogonalization and the QR-factorization.

1) Let $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ be the following vectors in R^3

$$\mathbf{X} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \qquad \mathbf{Y} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \qquad \mathbf{Z} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}.$$

a) Find $\mathbf{X} + \mathbf{Y}$ **b)** Find $\mathbf{Y} - \mathbf{Z}$ **c)** Find $\mathbf{X} - \mathbf{Y} + 2\mathbf{Z}$
d) Find a vector parallel to $\mathbf{Z}$ but of length two.

2) Find the angle between the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ -3 \\ -1 \end{pmatrix}.$$

3) Let $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ be the following vectors in R^3

$$\mathbf{X} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \qquad \mathbf{Y} = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}, \qquad \mathbf{Z} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

a) Find $\langle \mathbf{X}, \mathbf{Y} \rangle$, $||\mathbf{X}||$, $||\mathbf{Y}||$, and the cosine of the angle between $\mathbf{X}$ and $\mathbf{Y}$.
b) Find $\langle \mathbf{X}, \mathbf{Z} \rangle$ **c)** Find $\langle \mathbf{Y}, \mathbf{Z} \rangle$ **d)** Find $\langle 3\mathbf{X} - 100\mathbf{Y}, \mathbf{Z} \rangle$

4) Let L be the line through the points $P = (1, 3, -5)$ and $Q = (7, 5, -2)$. Find all vectors parallel to L having length one.

5) Given the vectors

$$\mathbf{x}_1 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 4 \\ 6 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 3 \\ -1 \\ -5 \end{pmatrix} \quad \mathbf{x}_4 = \begin{pmatrix} 6 \\ -2 \\ 3 \end{pmatrix},$$

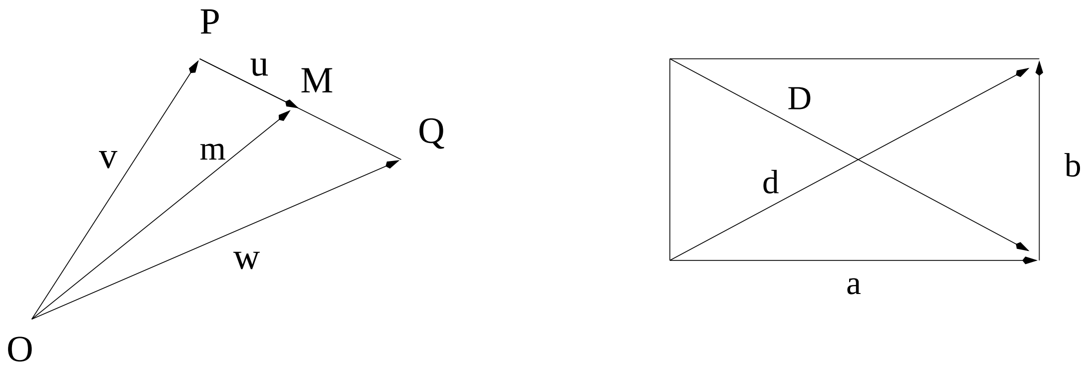

Figure 2: Figures for problems 6 and 8.

compute the quantities

$$\frac{\langle \mathbf{x}_1, \mathbf{x}_2 \rangle}{\langle \mathbf{x}_2, \mathbf{x}_2 \rangle} \cdot \mathbf{x}_2 \qquad \frac{\langle \mathbf{x}_4, \mathbf{x}_3 \rangle}{\langle \mathbf{x}_4, \mathbf{x}_4 \rangle} \cdot \mathbf{x}_4 \qquad \frac{1}{\langle \mathbf{x}_3, \mathbf{x}_3 \rangle} \cdot \mathbf{x}_3.$$

6) Let $\mathbf{v}$ and $\mathbf{w}$ be the tridimensional vectors represented in figure 1 above as arrow from the point O to the points P and Q, respectively. Let

$$\mathbf{m} = \frac{1}{2}(\mathbf{v} + \mathbf{w}).$$

Prove that the tip of the vector $\mathbf{m}$ is the midpoint of the line segment joining the tips of $\mathbf{v}$ and $\mathbf{w}$, as shown in Fig. 2.

7) Which angle is a right angle in the triangle with vertices $(0,0,0), (1,-1,2)$, and $(-4,0,5)$?

8) Prove this plane geometry theorem using vectors: The sum of squares of the sides of a parallelogram equals the sum of the squares of the diagonals. A parallelogram is depicted in figure 1, having sides represented by the vectors $\mathbf{a}$ and $\mathbf{b}$, and diagonals represented by the vectors $\mathbf{d}$ and $\mathbf{D}$.

9) Find the equation of the plane through the points $(1,2,-1), (2,3,1)$, and $(3,-1,2)$.

10) Consider a nonzero vector $\mathbf{u}$ represented by an arrow from point O to point P. Consider the vector $\mathbf{v}$ represented by an arrow from the point O to the point Q. The projection of $\mathbf{v}$ on $\mathbf{u}$ is defined to be the vector from O to N, where N is the foot of the perpendicular drawn from Q to the line of $\mathbf{u}$. Show that this projection is given by

$$\frac{\mathbf{u}\cdot\mathbf{v}}{|\ \mathbf{u}\ |^2}\mathbf{u}$$

and compute the projection of $(1,-2,1)$ on $(3,1,-4)$.

11) Consider a plane represented implicitly by $ax+by+cz=d$. Consider the vector $\mathbf{v}$ represented by an arrow from the point (x_0,y_0,z_0) in the plane to a point Q not in the plane. The projection of $\mathbf{v}$ on the plane is defined to be the vector from (x_0,y_0,z_0) to N, the perpendicular drawn from Q to the plane. Let $\mathbf{u}=(a,b,c)$. Show that the projection of $\mathbf{v}$ on the plane is given by

$$\mathbf{v}-\frac{\mathbf{u}\cdot\mathbf{v}}{|\ \mathbf{u}\ |^2}\cdot\mathbf{u},$$

and compute the projection of the vector $(2,3,-1)$ on the plane given by the equation $x-3y+2z=17$.

12) Show that the distance from point (x_1,y_1,z_1) to the plane represented by $ax+by+cz=d$ is given by

$$\frac{|\ ax_1+by_1+cz_1-d\ |}{\sqrt{a^2+b^2+c^2}}.$$

13) Given the vectors

$$\mathbf{X}=\begin{pmatrix}2\\-3\\0\end{pmatrix}\qquad\text{and}\qquad\mathbf{Y}=\begin{pmatrix}1\\-2\\3\end{pmatrix},$$

find the length of the projection of the vector $\mathbf{X}$ on $\mathbf{Y}$.

14) Prove that the length of the projection of the vector $\mathbf{x} = x_1\mathbf{e_1} + x_2\mathbf{e_2} + x_3\mathbf{e_3}$ onto the vector $\mathbf{a} = a_1\mathbf{e_1} + a_2\mathbf{e_2} + a_3\mathbf{e_3}$ is given by

$$\frac{a_1x_1 + a_2x_2 + a_3x_3}{\sqrt{a_1^1 + a_2^2 + a_3^2}}.$$

15) Find the cosine of the angle between an edge of a cube and a diagonal.

16) Use inner products to write the vector

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

as a linear combination of the vectors

$$\mathbf{q}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \mathbf{q}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \qquad \mathbf{q}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

17) Find values for a, b and c such that the following three vectors

$$\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} a \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} b \\ 3 \\ c \end{pmatrix}$$

are orthogonal.

18) Construct a set of orthonormal vectors from the set

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

19) Use the Gram-Schmidt orthogonalization process to generate an orthogonal basis for the subspace $\mathcal{W} = span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, where

$$\mathbf{u}_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

20) Find a basis for the null space of

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 1 & -5 \\ 2 & 1 & 7 & 5 \\ 1 & -1 & 2 & -2 \end{pmatrix}$$

and the use the Gram-Schmidt process to obtain an orthonormal basis.

21) Let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$, where

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_4 = \begin{pmatrix} 0 \\ 1 \\ -2 \\ 1 \end{pmatrix},$$

be a basis for R^4. Use the Gram-Schmidt process to transform the given basis into an orthonormal basis for R^4.

22) Let $\mathbf{a}$ be a fixed vector in R^3 and define $\mathcal{W}$ to be the subset of R^3 given by

$$\mathcal{W} = \{\mathbf{x} : \langle \mathbf{a}, \mathbf{x} \rangle = 0\}.$$

Prove that $\mathcal{W}$ is a subspace of R^3.

23) Let $\mathcal{U}$ and $\mathcal{V}$ be the subsets of R^3 defined by

$$\mathcal{U} = \{\mathbf{x} : \langle \mathbf{a}, \mathbf{x} \rangle = 0\}$$

and

$$\mathcal{V} = \{\mathbf{x} : \langle \mathbf{b}, \mathbf{x} \rangle = 0\},$$

where

$$\mathbf{a} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Prove that $\mathcal{U} + \mathcal{V}$ is NOT a subspace of R^3.

5.1 The QR-factorization.

Let $\mathbf{A}$ be an $m \times n$ matrix that has linearly independent column vectors $\{\mathbf{u}_1, \mathbf{u}_2, ...\mathbf{u}_n\}$. Suppose the Gram-Schmidt process has been applied to theses vectors in order to produce an orthonomal basis $\{\mathbf{q}_1, \mathbf{q}_2, ...\mathbf{q}_n\}$. If $\mathbf{Q}$ is the matrix $m \times n$ matrix with orthonormal column vectors $\{\mathbf{q}_1, \mathbf{q}_2, ...\mathbf{q}_n\}$ that results from applying the Gram-Schmidt process to the columns of $\mathbf{A}$, then $\mathbf{A}$ can be factored as $\mathbf{A}$=$\mathbf{QR}$, where $\mathbf{R}$ is an $n \times n$ invertible triangular matrix given by

$$\mathbf{R} = \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{q}_1 \rangle & \langle \mathbf{u}_2, \mathbf{q}_1 \rangle & ... & \langle \mathbf{u}_n, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{u}_2, \mathbf{q}_2 \rangle & ... & \langle \mathbf{u}_n, \mathbf{q}_2 \rangle \\ ... & ... & ... & ... \\ 0 & 0 & ... & \langle \mathbf{u}_n, \mathbf{q}_n \rangle \end{pmatrix}.$$

Note that $\mathbf{R}$ is invertible because its diagonal entries are all non-zero.

24) Find the **QR**-decomposition of

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

25) Determine the **QR** factors of

$$\begin{pmatrix} 0 & -20 & -14 \\ 3 & 27 & -4 \\ 4 & 11 & -2 \end{pmatrix}.$$

26) Construct the **QR**-decomposition for the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

27) Find the **QR**-factorization of

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 1 & 1 & -3 \\ 0 & 1 & 1 \end{pmatrix}.$$

5.2 Orthogonal matrices.

A square matrix with the property that

$$\mathbf{A}^{-1} = \mathbf{A}^T$$

is said to be an orthogonal matrix. For instance, the matrix

$$A = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{pmatrix}$$

is orthogonal.

28) The standard matrix for the counterclockwise rotation in R^2 through an angle θ is given by

$$\mathbf{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Show that this matrix is orthogonal for all choices of θ.

Given an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_k\}$, the matrix $\mathbf{Q} = [\mathbf{q}_1 \ \ \mathbf{q}_2 \ \ ... \ \ \mathbf{q}_k]$ having as columns the given basis vectors is an orthogonal matrix. For example, the matrix

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

is orthogonal since its column vectors form an orthonormal set in R^2.

29) Which of the following matrices are orthogonal?

a) $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{pmatrix}$ **b)** $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ **c)** $\frac{1}{13} \begin{pmatrix} 12 & 5 \\ -5 & 12 \end{pmatrix}$

d) $\begin{pmatrix} 2/3 & -2/3 & 1/3 \\ 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \end{pmatrix}$ **e**) $\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ **f**) $\begin{pmatrix} 2 & 0 \\ 1 & 1/2 \end{pmatrix}$.

30) Under what conditions on the real numbers α and β will the matrix

$$\mathbf{P} = \begin{pmatrix} \alpha+\beta & \beta-\alpha \\ \alpha-\beta & \beta+\alpha \end{pmatrix}$$

be an orthogonal matrix?

6 The problem of least squares – I.

Suppose that an experimenter collects data by taking the measurements $z_1, z_2, ..., z_n$ of certain physical quantity at certain times $t_1, t_2, ..., t_n$, respectively. Suppose that when the values of the variable z are plotted against time t it is apparent that there might be a linear relationship between z and t, as shown in Fig. 3. The pairs $(z_1, t_1), (z_2, t_2), ...(z_n, t_n)$ fall roughly along a straight line drawn in red. Let $z = a + bt$ be the equation of the straight line through the given points. For the times $t_1, t_2, ..., t_n$ the corresponding values of z on the curve are $\bar{z}_1, \bar{z}_2, ..., \bar{z}_n$, represented in the figure as red dots. In some intuitive sense, the actual value z_j that the experimenter has measured is relatively close to the corresponding values $\bar{z}_j$ given by the equation $z = a + bt$ for the same time t_j. We say then the the curve *fits* the given data.

Example 1. A soap manufacturer graphs its monthly records on soap sales versus total advertising expenditures and finds that these figures fall roughly along a straight line. Thus, if z_i equals the sales in the i^th month and t_i represents the advertising expenditures in the i^th month, and the records for a five-month period comprise the pairs $(z, t) = (2, 1), (4, 2), (7, 3), (9, 4), (11, 5)$, it clear that if we want to find a relationship between advertising expenditures and sales that is a linear function of the form $y = a + bt$, then the following equations must be satisfied simultaneously

$$\begin{aligned} a + b &= 2 \\ a + 2b &= 4 \\ a + 3b &= 7 \\ a + 4b &= 9 \\ a + 5b &= 11. \end{aligned} \tag{11}$$

In matrix notation the above system is in the standard form $\mathbf{Ax} = \mathbf{b}$,

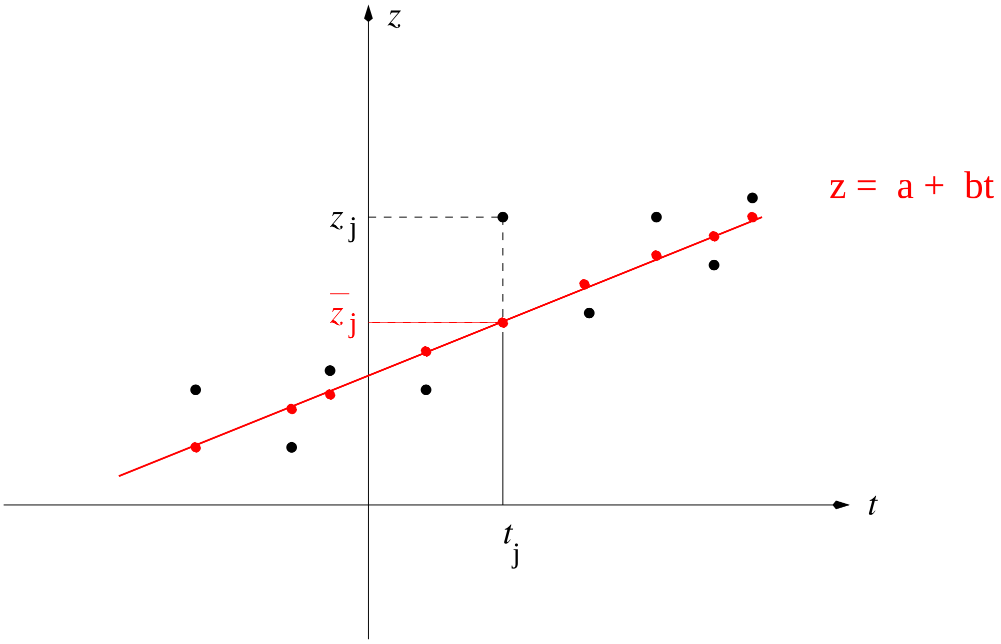

Figure 3: The least-squares fitting of data problem. The (red) straight line $z = a + bt$ is an approximation for the data points $(z_1, t_1), (z_2, t_2), \ldots (z_n, t_n)$. If the data points do not lie on a line, we seek the straight line that most closely fits the data. This best line is called the least-squares line of approximation. In statistics the fitting of a line is called linear regression and the line is called the linear regression line for the data.

where we define the matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 4 \\ 7 \\ 9 \\ 11 \end{pmatrix}.$$

As the above example suggests, the problem of fitting data translates into finding a solution to the system $\mathbf{Ax} = \mathbf{b}$ where, in a general context, we define

$$\mathbf{A} = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & t_n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_n \end{pmatrix}, \tag{12}$$

However, the above example also illustrate some important difficulties with the determination of the straight line that the fits the data. Note first that the system (11) has more equation than unknowns since the straight line $z = a + bt$ is completely determined by two parameters only, namely, the values of a and b. On the other hand, the data comprise 5 pair of values, and thus give rise to 5 equations to be satisfied by a and b.

When a system has more equations than unknowns it is called an overdetermined system. Overdetermined systems that arise in actual physical experiments as an empirical relationship between various quantities have, in general, no solution. The system (11) of *Example 1* is obviously inconsistent. This inconsistency means that it is not possible to determine a line that passes exactly through all the points of the data unless all the points <u>do</u> fall on a straight line. In this case, the vector $\mathbf{x}$ is the actual solution to the system $\mathbf{Ax} = \mathbf{b}$. Since this is rarely the case with actual measurements, we have to face the problem of determine a straight line that most closely fits the data. In this case, the vector $\mathbf{x}$ is the best approximation to the inconsistent system $\mathbf{Ax} = \mathbf{b}$.

The criterion for *closeness* may vary a lot. On possible criterion might be to consider the sum

$$|z_1 - \bar{z}_1| + |z_2 + \bar{z}_2| + ... + |z_n - \bar{z}_n|$$

and require that it be the minimum among all possible straight lines that can be drawn through the pair of points of the data. The line thus obtained would, in a certain sense, fit the data very well. There are, however, good reason not to choose this criterion.

The favored criterion is given by the so-called *least-squares method.* This method suggests that we consider the sum

$$(z_1 - \bar{z}_1)^2 + (z_2 + \bar{z}_2)^2 + ... + (z_n - \bar{z}_n)^2$$

and require that it be the minimum among all possible straight lines that can be drawn through the pair of points of the data.

To solve the least squares problem, write

$$\begin{aligned} F(a,b) &= (z_1 - \bar{z}_1)^2 + (z_2 - \bar{z}_2)^2 + ... + (z_n - \bar{z}_n)^2 \\ &= [z_1 - (a + bt_1)]^2 + [z_2 - (a + bt_2)]^2 + ... + [z_n - (a + bt_n)]^2 \end{aligned}$$

and minimize $F(a,b)$ as a function of a and b.

To do this, set the partial derivatives of F with respect to a and b equal to zero. The partial derivatives are given by

$$\begin{aligned} \frac{\partial F}{\partial a} &= -2[z_1 - (a + bt_1)] - ... - 2[z_n - (a + bt_n)] \\ &= -2\sum_{j=1}^{n} z_j + 2an + 2b\sum_{j=1}^{n} t_j \end{aligned}$$

and

$$\begin{aligned} \frac{\partial F}{\partial b} &= -2[z_1 - (a + bt_1)]t_1 - ... - 2[z_n - (a + bt_n)]t_n \\ &= -2\sum_{j=1}^{n} z_j t_j + 2b\sum_{j=1}^{n} t_j^2 + 2a\sum_{j=1}^{n} t_j. \end{aligned}$$

By setting $\partial F/\partial a = 0$ and $\partial F/\partial b = 0$ we obtain the following system of equations for a and b

$$\begin{aligned} na \quad &+ \quad (\sum_{j=1}^{n} t_j)\, b \quad = \quad \sum_{j=1}^{n} z_j \\ (\sum_{j=1}^{n} t_j)\, a \quad &+ \quad (\sum_{j=1}^{n} t_j^2)\, b \quad = \quad \sum_{j=1}^{n} z_j t_j. \end{aligned} \tag{13}$$

The line $z = a + bt$, with a and b determined by the two equations above, is called the *regression line* ot *least-squares regression line.*

Example 2. Find the regression line for the points $(1,0),(1,1),(2,2)$ and $(4,3)$.

Solution. Let $z = a + bt$ be the regression line. The equation for the constants a and b are (13), with $n = 4$. The various quantities appearing there are

$$t_1 + t_2 + t_3 + t_4 = 6 \qquad\qquad z_1 + z_2 + z_3 + z_4 = 8$$

$$t_1^2 + t_2^2 + t_3^2 + t_4^2 = 14 \qquad\qquad z_1t_1 + z_2t_2 + z_3t_3 + z_4t_4 = 17.$$

The system of equations (13) for a and b is

$$4a + 6b = 8 \qquad\qquad 6a + 14b = 17,$$

from which we get $a = 1/2$ and $b = 1$. Thus, $z = 1/2 + t$.

There are two matrix representations of the system of equations (13). The first one is just a rearrangement of that system in the following matrix notation

$$\begin{pmatrix} n & \sum_{j=1}^n t_j \\ \sum_{j=1}^n t_j & \sum_{j=1}^n t_j^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n z_j \\ \sum_{j=1}^n z_j t_j \end{pmatrix}. \tag{14}$$

Next, we use the matrices defined in (12)

$$\mathbf{A} = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & t_n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_n \end{pmatrix}, \tag{15}$$

and the system (14) becomes

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}. \tag{16}$$

The system (16) is called the **normal equations**, and the vector $\mathbf{x}$ that satisfies (16) is called the *least-squares* solution to the inconsistent system $\mathbf{A}\mathbf{x} = \mathbf{b}$ as defined in (15).
If the matrix $\mathbf{A}^T\mathbf{A}$ is non-singular, then the least-squares problem admits a unique solution given by

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}. \tag{17}$$

Example 3. A manufacturer keeps a record of sales per month (z) and advertising costs (t). For the first 4 months of 2000 she finds that they are running as follows $(z,t) = (2,1),(3,2),(5,3)$ and $(7,4)$. Find the regression line for these data.

Solution. Set

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 7 \end{pmatrix}.$$

Hence

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 10 \\ 10 & 30 \end{pmatrix}.$$

Thus,

$$(\mathbf{A}^T\mathbf{A})^{-1} = \frac{1}{10}\begin{pmatrix} 15 & -5 \\ -5 & 2 \end{pmatrix}.$$

Therefore

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} = \frac{1}{10}\begin{pmatrix} 15 & -5 \\ -5 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 17/10 \end{pmatrix}.$$

The regression line is thus $z = a + bt = 1.7t$.

1) A manufacturer keeps a record of sales (in thousands of units per month) and advertising costs (in thousands of dollars). For the first 5 months of 2000 he finds that they are running as follows (This is Example 5)

advertising costs (x)	1	2	3	4	5
sales (y)	2	4	7	9	11

Find the regression line for these data, that is, find $y = a + bx$. On the basis of this regression line predict the sales when the advertising expenditure are set at 10 thousand dollars.

2) Find the regression line for the following points.

x	-3	-2	-1	0	1	2
y	4	2	1	-1	-2	-3

3) Find the least-squares regression line for the data.

x	1	3	4	6	9	10
y	-3	2	3	7	12	14

4) Consider the experimental observations given in the following table

t	1	4	8	11
y	1	2	4	5

Find the least-squares fit to the data.

5) Find the least-squares regression line to the following data.

x	-3	-2	-1	0	1	2	3
s	10	15	19	27	28	34	42

6.1 The Problem of Least Squares – II

Formulas (16) and (17) have been applied to the system $\mathbf{Ax} = \mathbf{b}$ where the matrices $\mathbf{A}$ and $\mathbf{b}$ are as given by (15) in the case of the least-squares fitting of data. These same formulas, however, can be applied to any system of m linear algebraic equations in n unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\ ... \qquad ... \qquad & \quad ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= b_m, \end{aligned} \tag{18}$$

of the general form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is an arbitrary $m \times n$ matrix and $\mathbf{b}$ is an $n \times 1$ column vector.
When this system is inconsistent there is no set of variables $x_1, x_2, ..., x_n$ that solve (18). As before, then, we want to find a set of variables that *most closely* solve the problem. This is the least-squares solution to $\mathbf{Ax} = \mathbf{b}$. It is again given by the normal equations (16).
In this general case, we will make the assumptions that the system $\mathbf{Ax} = \mathbf{b}$ is overdetermined, so that $m \geq n$, and that the matrix $\mathbf{A}$ has rank n. Under this assumptions the matrix $\mathbf{A}^T\mathbf{A}$ is $n \times n$, non-singular and invertible. Thus, the normal equations (16) admit a unique

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b},$$

as in (17).
The expression $\mathbf{A}^- = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ appears so often in any non-trivial matrix problem, that it is called the *generalized inverse* or the *Moore-Penrose* inverse of the matrix $\mathbf{A}$. Note that the matrix $(\mathbf{A}^T\mathbf{A})^{-1}$ is an $n \times n$ matrix and $\mathbf{A}^-$ is an $n \times m$ matrix.

Example 4. Find the least-squares solution to

$$\begin{aligned} x + 2y + z &= 1 \\ 3x - y &= 2 \\ 2x + y - z &= 2 \\ x + 2y + 2z &= 1 \end{aligned}$$

Solution 1. [*Direct application of (17)*] The above system of equations takes the form $\mathbf{A}\mathbf{x} = \mathbf{b}$ with

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & -1 & 0 \\ 2 & 1 & -1 \\ 1 & 2 & 2 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}.$$

Then

$$\begin{aligned} \mathbf{x} &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \\ &= \begin{pmatrix} 15 & 3 & 1 \\ 3 & 10 & 5 \\ 1 & 5 & 6 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 3 & 2 & 1 \\ 2 & -1 & 1 & 2 \\ 1 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.7597 \\ 0.2607 \\ -0.1772 \end{pmatrix}. \end{aligned}$$

Solution 2. [*Using the normal equations*] In this solution we illustrate the procedure to follow when one does not want to invert the 3×3 matrix $\mathbf{A}^T\mathbf{A}$. The alternative is to solve the normal equations. The above system of equations takes the form $\mathbf{A}\mathbf{x} = \mathbf{b}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & -1 & 0 \\ 2 & 1 & -1 \\ 1 & 2 & 2 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix},$$

then

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 15 & 3 & 1 \\ 3 & 10 & 5 \\ 1 & 5 & 6 \end{pmatrix} \quad \text{and} \quad \mathbf{A}^T\mathbf{b} = \begin{pmatrix} 12 \\ 4 \\ 1 \end{pmatrix},$$

and the normal equation becomes

$$\begin{pmatrix} 15 & 3 & 1 \\ 3 & 10 & 5 \\ 1 & 5 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 12 \\ 4 \\ 1 \end{pmatrix},$$

from which we obtain $x = 0.7597$, $y = 0.2607$ and $z = -0.1772$.
Note that this is the only procedure available when the matrix $\mathbf{A}^T\mathbf{A}$ is singular.

6) Find the least-squares solution to

$$\begin{aligned} y &= 80 \\ 2x + 5y &= 100 \\ 5x - 2y &= 60 \\ -x + 8y &= 130 \\ 10x - y &= 150 \end{aligned}$$

7) Given

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix},$$

show that the system $\mathbf{Ax} = \mathbf{b}$ is inconsistent. Then find the least-squares solution of this inconsistent system.

8) Show that if $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{A}^{-}\mathbf{A} = \mathbf{I}_n$.

9) Show that if $\mathbf{A}$ is an $n \times n$ non-singular matrix, then $\mathbf{A}^{-} = \mathbf{A}^{-1}$.

10) Find a least-squares solution to the inconsistent system $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}.$$

11) Find a least-squares solution to $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 5 \\ -3 \\ -5 \end{pmatrix}.$$

Explain how the rank of $\mathbf{A}$ is related to your solution.

12) Find a least-squares solution to $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{pmatrix}.$$

From the solution you obtained what can you say about the rank of $\mathbf{A}$?

13) Show that if $\mathbf{A} = \mathbf{QR}$ is a QR-decomposition (Section 5) of the $m \times n$ matrix $\mathbf{A}$, the normal equation (16) can be written as

$$\mathbf{R}^T\mathbf{Rx} = \mathbf{R}^T\mathbf{Q}^T\mathbf{x}.$$

Prove that if $\mathbf{A}$ has linearly independent columns, the above equation has a unique solution

$$\mathbf{x} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{x}$$

if the columns of $\mathbf{A}$ are linearly independent.

14) Find the least-squares solution of $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 3 \\ 5 \\ 7 \\ -3 \end{pmatrix}$$

by using the QR-factorization of $\mathbf{A}$.

6.2 The Problem of Least Squares – III

The least-squares problem can also be stated as follows. Let $\mathcal{V}$ be a p-dimensional subspace of R^n. Given a vector $\mathbf{b}$ in R^n, find a vector $\mathbf{v}_0$ in $\mathcal{V}$ such that

$$||\mathbf{b} - \mathbf{v}_0|| \leq ||\mathbf{b} - \mathbf{v}|| \tag{19}$$

for all $\mathbf{v}$ in $\mathcal{V}$. The vector $\mathbf{v}_0$ is called the *orthogonal projection of the vector* $\mathbf{b}$ *on* $\mathcal{V}$ or *best least-squares approximation* to $\mathbf{b}$.

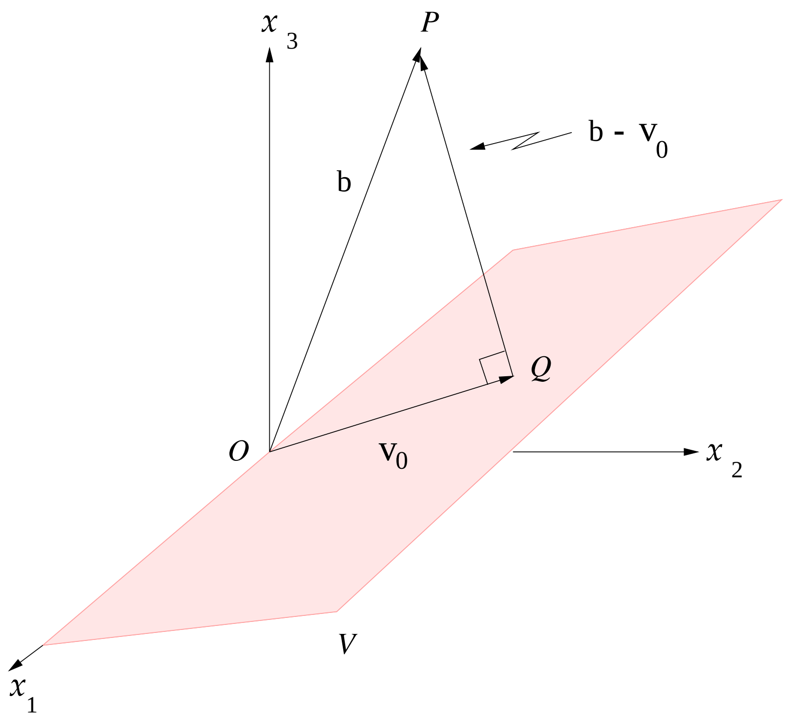

Figure 4: The least squares problem in three dimensions. The projection of $\mathbf{b}$ onto the subspace $\mathcal{V}$ is the vector $\mathbf{v}_0$ that minimizes the distance $||\mathbf{b} - \mathbf{v}||$ for all $\mathbf{v}$ in $\mathbf{V}$.

It is easy to understand the meaning of the least squares problem in three dimensions. Let $\mathcal{V}$ be a two-dimensional plane through the origin O, as depicted in Fig. 4. An arbitrary vector $\mathbf{b}$ of R^3 is shown from the origin to the point P. We want to find a special vector $\mathbf{v}_0$ in $\mathcal{V}$ such that $||\mathbf{b} - \mathbf{v}_0||$ is least among all the vectors in $\mathcal{V}$. This means that $\mathbf{v}_0$ is that vector in the plane that is closest to $\mathbf{b}$.

It is equally easy to find the solution to this problem in three dimensions. If $\mathbf{v}_0$ is represented by a vector from the origin to the point Q in the plane, then $\mathbf{b} - \mathbf{v}_0$ is represented by the vector from Q to P. The point Q that is closest to P is obtained by dropping a perpendicular from P to the plane $\mathcal{V}$ through Q.

In this case, the vector $\mathbf{v}_0$ will be the orthogonal projection of the vector $\mathbf{b}$ onto the plane $\mathcal{V}$, and the minimum distance between P and $\mathcal{V}$ will be given by

$$||\mathbf{b} - \mathbf{v}_0||.$$

This intuitive idea that the vector $\mathbf{v}_0$ that satisfies (19) is such that the difference $\mathbf{b} - \mathbf{v}_0$ is perpendicular to the plane suggests the following theorem in which this idea is made precise in higher dimensional spaces.

Theorem. Let $\mathcal{V}$ be a p-dimensional subspace of R^n and let $\mathbf{b}$ be a vector in R^n. Suppose that there is a vector $\mathbf{v}_0$ in $\mathcal{V}$ such that

$$\langle \mathbf{b} - \mathbf{v}_0, \mathbf{v} \rangle = 0$$

for every vector $\mathbf{v}$ in $\mathcal{V}$. Then for any $\mathbf{v}$ in $\mathcal{V}$,

$$||\mathbf{b} - \mathbf{v}_0|| \leq ||\mathbf{b} - \mathbf{v}||.$$

Proof. Let $\mathbf{v}$ be any vector in $\mathcal{V}$. We can write

$$||\mathbf{b} - \mathbf{v}|| = ||(\mathbf{b} - \mathbf{v}_0) + (\mathbf{v}_0 - \mathbf{v})||. \tag{20}$$

Note that $\mathbf{v}_0 - \mathbf{v}$, being a difference of vectors in $\mathcal{V}$, is in $\mathcal{V}$, and that $\mathbf{b} - \mathbf{v}_0$ is orthogonal to $\mathcal{V}$, so that the two terms on the right side of (20) are orthogonal. Thus, by the Theorem of Pythagoras, we obtain

$$||\mathbf{b} - \mathbf{v}||^2 = ||\mathbf{b} - \mathbf{v}_0||^2 + ||\mathbf{v}_0 - \mathbf{v}||^2.$$

But $||\mathbf{v}_0 - \mathbf{v}||^2 \geq 0$, and it follows that

$$||\mathbf{b} - \mathbf{v}|| \geq ||\mathbf{b} - \mathbf{v}_0||.$$

The fact that $\mathbf{v}_0$ is the best approximation to $\mathbf{b}$ whenever $\mathbf{b} - \mathbf{v}_0$ is perpendicular to the plane in the three dimensional case yields a procedure to find the best approximation in higher dimensions. This procedure relies on the observation that if the vector $\mathbf{b} - \mathbf{v}_0$ is perpendicular to $\mathcal{V}$, it is perpendicular to all vectors in $\mathcal{V}$. Thus, if $\mathcal{V}$ is a p-dimensional subspace with basis vector

$\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p\}$, the vector $\mathbf{b} - \mathbf{v}_0$ is perpendicular to all the p vectors of this basis we have

$$\begin{aligned}
\langle \mathbf{b} - \mathbf{v}_0,\ \mathbf{v}_1 \rangle &= 0 \\
\langle \mathbf{b} - \mathbf{v}_0,\ \mathbf{v}_2 \rangle &= 0 \\
\dots \quad \dots \quad &\quad \dots \\
\langle \mathbf{b} - \mathbf{v}_0,\ \mathbf{v}_p \rangle &= 0.
\end{aligned} \tag{21}$$

The set of equations (21) determines the orthogonal projection of $\mathbf{b}$ onto $\mathcal{V}$.

Example 5. Let $\mathcal{V}$ be the subspace of R^3 defined by

$$\mathbf{V} = \{\mathbf{v} : \mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, x_1 + x_2 - 3x_3 = 0\}.$$

Let $\mathbf{b}$ be the vector

$$\mathbf{b} = \begin{pmatrix} 1 \\ -2 \\ -4 \end{pmatrix}.$$

Find the orthogonal projection of the vector $\mathbf{b}$ onto $\mathcal{V}$.

Solution. We first determine a basis for $\mathbf{V}$. One possible basis is formed by the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

The orthogonal projection of the vector $\mathbf{b}$ on $\mathcal{V}$ is a vector $\mathbf{v}_0$ that satisfies the equation

$$\langle \mathbf{b} - \mathbf{v}_0, \mathbf{v}_1 \rangle = 0 \quad \text{and} \quad \langle \mathbf{b} - \mathbf{v}_0, \mathbf{v}_2 \rangle = 0, \tag{22}$$

since in this case there are only two equations in the set (21). Since the orthogonal projection $\mathbf{v}_0$ is a vector in $\mathcal{V}$, it can be expressed in terms of the basis vectors as

$$\mathbf{v}_0 = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2, \tag{23}$$

where the constants x_1 and x_2 can be determined upon substitution of (23) into the set (22). We get

$$2x_1 + 3x_2 = 3 \qquad \text{and} \qquad 3x_1 + 10x_2 = -1,$$

from which we obtain $x_1 = 3$ and $x_2 = -1$. Thus,

$$\mathbf{v}_0 = 3\mathbf{v}_1 - \mathbf{v}_2 = \begin{pmatrix} 0 \\ -3 \\ -1 \end{pmatrix}.$$

Let us now determine $\mathbf{v}_0$ in the general setting of equations (21). If $\mathcal{V}$ is p-dimensional with basis vectors $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p\}$, we can always write

$$\mathbf{v}_0 = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_p\mathbf{v}_p. \tag{24}$$

Substituting this expression for $\mathbf{v}_0$ into (21) we obtain

$$\begin{aligned} \langle \mathbf{b} - (x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_p\mathbf{v}_p),\ \mathbf{v}_1 \rangle &= 0 \\ \langle \mathbf{b} - (x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_p\mathbf{v}_p),\ \mathbf{v}_2 \rangle &= 0 \\ ... \qquad ... \qquad &... \\ \langle \mathbf{b} - (x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_p\mathbf{v}_p),\ \mathbf{v}_p \rangle &= 0. \end{aligned}$$

or, with some minor changes,

$$\begin{aligned} \langle \mathbf{v}_1,\ \ x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_p\mathbf{v}_p \rangle &= \langle \mathbf{v}_1,\ \mathbf{b} \rangle \\ \langle \mathbf{v}_2,\ \ x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_p\mathbf{v}_p \rangle &= \langle \mathbf{v}_2,\ \mathbf{b} \rangle \\ ... \qquad ... \qquad &\quad ... \\ \langle \mathbf{v}_p,\ \ x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_p\mathbf{v}_p \rangle &= \langle \mathbf{v}_p,\ \mathbf{b} \rangle \end{aligned} \tag{25}$$

By distributing the inner products in (25) the system takes the form

$$\begin{aligned} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle x_1 + \langle \mathbf{v}_1, \mathbf{v}_2 \rangle x_2 + ... + \langle \mathbf{v}_1, \mathbf{v}_p \rangle x_p &= \langle \mathbf{v}_1, \mathbf{b} \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle x_1 + \langle \mathbf{v}_2, \mathbf{v}_2 \rangle x_2 + ... + \langle \mathbf{v}_2, \mathbf{v}_p \rangle x_p &= \langle \mathbf{v}_2, \mathbf{b} \rangle \\ ... \qquad ... \qquad &\quad ... \\ \langle \mathbf{v}_p, \mathbf{v}_1 \rangle x_1 + \langle \mathbf{v}_p, \mathbf{v}_2 \rangle x_2 + ... + \langle \mathbf{v}_p, \mathbf{v}_p \rangle x_p &= \langle \mathbf{v}_p, \mathbf{b} \rangle. \end{aligned} \tag{26}$$

We can write this system of equation even more compactly if we define $\mathbf{A}$ to be the $n \times p$ matrix whose columns are the basis vectors of the subspace, $\mathbf{A} = [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \ldots \;\; \mathbf{v}_p]$. Then (26) becomes

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}, \tag{27}$$

where $\mathbf{x}$ is the column vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_p \end{pmatrix}.$$

Thus, the problem of finding orthogonal projections has been converted to a problem of finding an exact solution to the associated normal system identical to (16). In this case, the $n \times p$ matrix $\mathbf{A}$ has linearly independent column vectors. This guarantees that the matrix $\mathbf{A}^T\mathbf{A}$ is non-singular and invertible. Thus, the vector $\mathbf{x}$ is given by

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}. \tag{28}$$

In this notation, the orthogonal projection vector (24) is written as

$$\mathbf{v}_0 = \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} = \mathbf{P}_v\mathbf{b}, \tag{29}$$

where we have defined $\mathbf{P}_v$ to be the $p \times p$ matrix

$$\mathbf{P}_v = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$$

called the *projection matrix* for $\mathcal{V}$.

Example 6. Find the projection matrix for the subspace $\mathcal{V}$ defined by

$$\mathbf{v} = \{\mathbf{v} : \mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \; x_1 + x_2 - 3x_3 = 0\},$$

and use it to find $\mathbf{P}_v\mathbf{b}$ for the vector $\mathbf{b}$ of *Example 5.*

Solution. This is the same subspace of Example 5. We need to choose <u>any</u> basis for this susbspace. Let us choose a basis that is different from the one we used in Example 5. We pick

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Construct the matrix $\mathbf{A} = [\mathbf{v}_1 \;\; \mathbf{v}_2]$,

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{with} \quad \mathbf{A}^T = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

Form the product

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 6 & 5 \\ 5 & 6 \end{pmatrix} \quad \text{with} \quad (\mathbf{A}^T\mathbf{A})^{-1} = \frac{1}{11}\begin{pmatrix} 6 & -5 \\ -5 & 6 \end{pmatrix}.$$

Therefore,

$$\mathbf{P}_v = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \frac{1}{11}\begin{pmatrix} 10 & -1 & 3 \\ -1 & 10 & 3 \\ 3 & 3 & 2 \end{pmatrix}.$$

The orthogonal projection of the vector $\mathbf{b} = \begin{pmatrix} 1 \\ -2 \\ -4 \end{pmatrix}$ onto $\mathcal{V}$ is given by

$$\mathbf{P}_v\mathbf{b} = \frac{1}{11}\begin{pmatrix} 10 & -1 & 3 \\ -1 & 10 & 3 \\ 3 & 3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \\ -4 \end{pmatrix} = \begin{pmatrix} 0 \\ -3 \\ -1 \end{pmatrix},$$

exactly as in Example 5.

15) Find $\mathbf{P}_v^2$ for the projection matrix of *Example 6.*

Example 7. Find the orthogonal projection of the vector

$$\mathbf{b} = \begin{pmatrix} -3 \\ -3 \\ 8 \\ 9 \end{pmatrix}$$

on the subspace of R^4 spanned by the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 2 \\ -1 \end{pmatrix}.$$

Solution. The subspace $\mathcal{V}$ of R^4 spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is the column space of the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{pmatrix}.$$

We have

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 11 & 6 & -4 \\ 6 & 7 & 0 \\ -4 & 0 & 6 \end{pmatrix} \quad \text{and} \quad (\mathbf{A}^T\mathbf{A})^{-1} = \frac{1}{134}\begin{pmatrix} 42 & -36 & 28 \\ -36 & 50 & -24 \\ 28 & -24 & 41 \end{pmatrix}.$$

Therefore,

$$\begin{aligned} \mathbf{v}_0 &= \begin{pmatrix} 3 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{pmatrix} \frac{1}{134}\begin{pmatrix} 42 & -36 & 28 \\ -36 & 50 & -24 \\ 28 & -24 & 41 \end{pmatrix}\begin{pmatrix} 3 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ -1 & 0 & 2 & -1 \end{pmatrix}\begin{pmatrix} -3 \\ -3 \\ 8 \\ 9 \end{pmatrix} \\ &= \begin{pmatrix} -2 \\ 3 \\ 4 \\ 0 \end{pmatrix}. \end{aligned}$$

16) Consider the matrix $\mathbf{M}$ and the vector $\mathbf{b}$, where

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \\ -1 & 1 & -1 \\ -1 & 2 & 0 \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} 3 \\ -3 \\ 0 \\ -3 \end{pmatrix}.$$

a) Find a basis for the columns space of the matrix $\mathbf{M}$.

b) Find the projection of the vector $\mathbf{b}$ onto the column space of $\mathbf{M}$.

c) Find $\mathbf{P}_v^2$.

17) Consider the matrix $\mathbf{M}$ and the vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, where

$$\mathbf{M} = \begin{pmatrix} 1 & 2 & 4 \\ -1 & 0 & -2 \\ 1 & 1 & 3 \end{pmatrix} \qquad \mathbf{b}_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

a) Find a basis for the columns space of the matrix $\mathbf{M}$.

b) Find the orthogonal projection of the vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ onto the column space of $\mathbf{M}$.

18) Find a basis for the column space of the matrix

$$\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Find the projection matrix $\mathbf{P}_v$ and the projection of the vector $\mathbf{b} = \begin{pmatrix} 6 \\ 2 \\ 1 \end{pmatrix}$ onto the colum space of $\mathbf{B}$.

19) Find a basis for the subspace

$$\mathcal{W} = \{\mathbf{x} : \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad x_1 + x_2 + x_3 = 0, \; x_1 - x_2 - x_3 = 0\}.$$

For the vector

$$\mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix},$$

solve the normal equations and determine the projection of $\mathbf{b}$ onto the subspace $\mathcal{W}$.

20) Prove that the projection matrix $\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ has the following two basic properties

a) $\mathbf{P}^2 = \mathbf{P}$

b) $\mathbf{P}^T = \mathbf{P}$

7 Eigenvalues and eigenvectors.

1) Show that the characteristic equation for the eigenvalue problem $(A - \lambda I)\mathbf{x} = 0$, where

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

is $\lambda^3 - I_1\lambda^2 + I_2\lambda - I_3 = 0$, where $I_1 = a_{11} + a_{22} + a_{33}$ and

$$I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix},$$

and

$$I_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

2) Find the characteristic equation, the eigenvalues and the eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{pmatrix}.$$

3) Find the eigenvectors and eigenvalues of the matrix

$$\mathbf{R} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

that represents a rotation in the plane through an angle α. For which values of α are the eigenvalues imaginary?

4) Find bases for the eigenspace of

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

5) Prove that if $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \lambda_2, ..., \lambda_k$, then $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ is a linearly independent set.

6) Find the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

7) The Hamilton-Cayley theorem asserts that every square matrix satisfies its own characteristic equation. That is, if

$$det(\mathbf{A} - \lambda\mathbf{I}) = \alpha_n\lambda^n + \alpha_{n-1}\lambda^{n-1} + ... + \alpha_2\lambda^2 + \alpha_1\lambda + \alpha_0 = 0,$$

then

$$\alpha_n\mathbf{A}^n + \alpha_{n-1}\mathbf{A}^{n-1} + ... + \alpha_2\mathbf{A}^2 + \alpha_1\mathbf{A} + \alpha_0\mathbf{I} = 0.$$

8) Verify the Hamilton-Cayley theorem for the matrix

$$\begin{pmatrix} 3 & 5 \\ -2 & -4 \end{pmatrix}.$$

9) Determine the eigenvalues and eigenvectors of

$$\mathbf{A} = \begin{pmatrix} -3 & 1 & -3 \\ 20 & 3 & 10 \\ 2 & -2 & 4 \end{pmatrix}.$$

10) Show that **i)** the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

has precisely **three** linearly independent eigenvectors; **ii)** the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

has precisely **two** linearly independent eigenvectors; **iii)** the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

has precisely **one** linearly independent eigenvector.

11) Use the Hamilton-Cayley theorem to find the inverse of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & -3 \\ 2 & 2 & 4 \end{pmatrix}.$$

12) Find the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix}.$$

13) Find the characteristic polynomial of the matrix

$$\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

14) Determine the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

15) Find an orthonormal set of eigenvectors for

$$\mathbf{A} = \begin{pmatrix} 7 & -16 & -8 \\ -16 & 7 & 8 \\ -8 & 8 & -5 \end{pmatrix}.$$

16) Use the Hamilton-Cayley theorem to compute $\mathbf{A}^8 - 3\mathbf{A}^5 + \mathbf{A}^4 + \mathbf{A}^2 - 4\mathbf{I}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

17) Show that the characteristic equation of the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{pmatrix}$$

is $\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$.

18) Determine the characteristic roots and eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 4 & -3 & -2 \\ 2 & -1 & -2 \\ 3 & -3 & -1 \end{pmatrix}.$$

Find a matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal.

19) Determine the characteristic roots and eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 1 & -3 \end{pmatrix}.$$

Find a matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal.

20) Find the characterisitc polynomial of the <u>block triangular</u> matrix

$$\begin{pmatrix} 9 & -1 & 5 & 7 \\ 8 & 3 & 2 & -4 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & -1 & 8 \end{pmatrix}.$$

8 Linear Transformations and Similarity.

8.1 Coordinate systems and bases

Let $\mathcal{V}$ be a k-dimensional subspace of R^n and $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ be a basis for $\mathcal{V}$.
Each vector $\mathbf{v}$ in V can be uniquely expressed as a linear combination of the basis vectors as

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_k\mathbf{v}_k, \tag{30}$$

where the coefficients $x_1, x_2, ..., x_k$ are called the coordinates of the vector $\mathbf{v}$ with respect to the basis $\mathcal{B}_V$.
In matrix notation the above equation reads

$$\mathbf{v} = \begin{pmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_k \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_k \end{pmatrix} = \mathbf{V} \cdot \mathbf{x}_V, \tag{31}$$

where $\mathbf{V} = [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \dots \;\; \mathbf{v}_k]$ is a $n \times k$ matrix having as columns the basis vectors and $\mathbf{x}_V$ is the coordinate vector of $\mathbf{v}$ relative to the basis $\mathcal{B}_V$.

Example 1. Let $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a basis for R^3, where

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 2 \\ -3 \end{pmatrix}.$$

The vector

$$\mathbf{v} = \begin{pmatrix} -3 \\ 1 \\ 5 \end{pmatrix}$$

can be written as a linear combination of the basis vectors in the following way.

$$\begin{aligned}
\mathbf{v} &= \frac{1}{7}\begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} + \frac{13}{7}\begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix} + \frac{1}{7}\begin{pmatrix} 3 \\ 2 \\ -3 \end{pmatrix} \\
&= \begin{pmatrix} 2 & -2 & 3 \\ 5 & 0 & 2 \\ -1 & 3 & -3 \end{pmatrix}\begin{pmatrix} 1/7 \\ 13/7 \\ 1/7 \end{pmatrix} \\
&= \mathbf{V}\mathbf{x}_V.
\end{aligned}$$

Note that if the $\mathcal{V} = R^n$, then $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_{k=n}]$ is a square matrix, since $k = n$, and invertible, since its columns are linearly independent vectors. In this case, $\mathbf{V}^{-1}$ exists and we can use the equation $\mathbf{v} = \mathbf{V}\mathbf{x}_V$ to obtain the coordinates of a vector $\mathbf{v}$ relative to a basis $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in the following way

$$\mathbf{x}_V = \mathbf{V}^{-1}\mathbf{v}. \tag{32}$$

Example 2. Use formula (32) to find the coordinates of the vector

$$\mathbf{v} = \begin{pmatrix} -3 \\ 1 \\ 5 \end{pmatrix}$$

relative to the basis $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 2 \\ -3 \end{pmatrix}.$$

Solution. We have

$$\begin{aligned}
\mathbf{x}_V &= \mathbf{V}^{-1}\mathbf{v} \\
&= \begin{pmatrix} 2 & -2 & 3 \\ 5 & 0 & 2 \\ -1 & 3 & -3 \end{pmatrix}^{-1}\begin{pmatrix} -3 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 1/7 \\ 13/7 \\ 1/7 \end{pmatrix}.
\end{aligned}$$

Suppose that a second basis $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_k\}$ is chosen for the subspace $\mathcal{V}$ of R^n. Each vector $\mathbf{v}$ in $\mathcal{V}$ can be alternatively written as a linear combination of the basis vectors in $\mathcal{B}_W$. Thus,

$$\mathbf{v} = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + ... + y_n\mathbf{w}_k, \tag{33}$$

or, in matrix notation,

$$\mathbf{v} = \begin{pmatrix} | & | & & | \\ \mathbf{w}_1 & \mathbf{w}_2 & ... & \mathbf{w}_k \\ | & | & & | \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_k \end{pmatrix} = \mathbf{W} \cdot \mathbf{x}_W, \tag{34}$$

where $\mathbf{W} = [\mathbf{w}_1 \;\; \mathbf{w}_2 \;\; ... \;\; \mathbf{w}_k]$ is a $n \times k$ matrix having as columns the new basis vectors and $\mathbf{x}_W$ is the coordinate vector of $\mathbf{v}$ relative to the basis $\mathcal{B}_W$.

Example 3. Let $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ be a basis for R^3, where

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix}, \qquad \mathbf{w}_3 = \begin{pmatrix} 8 \\ 1 \\ -10 \end{pmatrix}.$$

The vector

$$\mathbf{v} = \begin{pmatrix} -3 \\ 1 \\ 5 \end{pmatrix}$$

can be written as a linear combination of the basis vectors in the following way.

$$\begin{aligned} \mathbf{v} &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{5}\begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix} - \frac{2}{5}\begin{pmatrix} 8 \\ 1 \\ -10 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 4 & 8 \\ 1 & -2 & 1 \\ 1 & 0 & -10 \end{pmatrix} \begin{pmatrix} 1 \\ -1/5 \\ -2/5 \end{pmatrix} \\ &= \mathbf{W}\mathbf{x}_W. \end{aligned} \tag{35}$$

Since any vector $\mathbf{v}$ in $\mathcal{V}$ can be expressed equally well as a linear combination of the basis vectors of $\mathcal{B}_V$ as in (31), or as a linear combination of the basis vectors of $\mathcal{B}_W$ as (34), we can establish a connection between the coordinate vectors of $\mathbf{v}$ relative to the two bases through the following equality.

$$\mathbf{v} = \mathbf{V}\mathbf{x}_V = \mathbf{W}\mathbf{x}_W. \tag{36}$$

Example 4. Suppose you know the coordinates of the vector $\mathbf{v}$ relative to the basis $\mathcal{B}_V$ given in Example 1. Find the coordinates of that same vector with respect to the basis $\mathcal{B}_W$ given in Example 3.

Solution. Since $\mathbf{W}$ is a 3×3 non-singular matrix we can use (13) in the following way.

$$\begin{aligned}
\mathbf{x}_W &= \mathbf{W}^{-1}\mathbf{V}\mathbf{x}_V \\
&= \begin{pmatrix} 1 & 4 & 8 \\ 1 & -2 & 1 \\ 1 & 0 & -10 \end{pmatrix}^{-1} \begin{pmatrix} 2 & -2 & 3 \\ 5 & 0 & 2 \\ -1 & 3 & -3 \end{pmatrix} \begin{pmatrix} 1/7 \\ 13/7 \\ 1/5 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ -1/5 \\ -2/5 \end{pmatrix},
\end{aligned}$$

as we have already determined in example 3.

8.2 The transition matrix

We must stress that in general both $\mathbf{V}$ and $\mathbf{W}$ are $n \times k$ matrices, not necessarily square matrices, so that we cannot invert $\mathbf{W}$ as we did in the example above and also in Example 2. In this case, we need to be more general in order to obtain a connection between coordinate vectors relative to two different bases. To do so, we begin by writing each vector of $\mathcal{B}_W$ as a linear combination of the vectors of $\mathcal{B}_V$. We have

$$\begin{aligned}
\mathbf{w}_1 &= s_{11}\mathbf{v}_1 + s_{21}\mathbf{v}_2 + \ldots + s_{k1}\mathbf{v}_k \\
\mathbf{w}_2 &= s_{12}\mathbf{v}_1 + s_{22}\mathbf{v}_2 + \ldots + s_{k2}\mathbf{v}_k \\
\ldots & \quad\;\; \ldots \qquad\qquad \ldots \\
\mathbf{w}_k &= s_{1k}\mathbf{v}_1 + s_{2k}\mathbf{v}_2 + \ldots + s_{kk}\mathbf{v}_k
\end{aligned} \tag{37}$$

Each of the $\mathbf{w}_j's$ above can be put in matrix notation analogous to that used in (31). Thus we also have

$$\mathbf{w}_1 = \begin{pmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_k \\ | & | & & | \end{pmatrix} \begin{pmatrix} s_{11} \\ s_{21} \\ \cdot \\ \cdot \\ \cdot \\ s_{k1} \end{pmatrix} = \mathbf{V} \cdot \mathbf{s}_1$$

$$\ldots \qquad \ldots \qquad \ldots \qquad \ldots$$

$$\mathbf{w}_k = \begin{pmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_k \\ | & | & & | \end{pmatrix} \begin{pmatrix} s_{1k} \\ s_{2k} \\ \cdot \\ \cdot \\ \cdot \\ s_{kk} \end{pmatrix} = \mathbf{V} \cdot \mathbf{s}_k,$$

where each $\mathbf{s}_j$ is the coordinate of the vector $\mathbf{w}_j$ relative to the basis $\mathcal{B}_V$. If we define a $k \times k$ matrix of the coefficients $\mathbf{S} = [\mathbf{s}_1 \ \ \mathbf{s}_2 \ \ \ldots \ \ \mathbf{s}_k]$, then we have

$$\begin{aligned} \mathbf{W} &= [\mathbf{w}_1 \ \ \mathbf{w}_2 \ \ \ldots \ \ \mathbf{w}_k] \\ &= [\mathbf{V}\mathbf{s}_1 \ \ \mathbf{V}\mathbf{s}_2 \ \ \ldots \ \ \mathbf{V}\mathbf{s}_k] \\ &= \mathbf{V}[\mathbf{s}_1 \ \ \mathbf{s}_2 \ \ \ldots \ \ \mathbf{s}_k] \\ &= \mathbf{V}\mathbf{S} \end{aligned} \tag{38}$$

Note that while $\mathbf{V}$ and $\mathbf{W}$ are $n \times k$ matrices, the matrix $\mathbf{S}$ is $k \times k$. Also it is not difficult to see that $\mathbf{S}$ is invertible. If it is, then the equation $\mathbf{S}\mathbf{x} = \mathbf{0}$ has only the trivial solution. But then $\mathbf{V}\mathbf{S}\mathbf{x} = \mathbf{W}\mathbf{x} = 0$ as well.

The columns of $\mathbf{W}$ are linearly independent, and thus $\mathbf{W}\mathbf{x} = \mathbf{0}$ admits only the trivial solution $\mathbf{x} = 0$. As a result, $\mathbf{S}\mathbf{x} = 0$ has only the trivial solution. This implies that the columns of $\mathbf{S}$ are linearly independent. Since $\mathbf{S}$ is a square matrix, it follows that it is invertible.

Example 5. Verify the formula (38) for the bases $\mathcal{B}_V$ and $\mathcal{B}_W$ of Examples 1 and 3.

Solution. We first write the basis vectors $\mathcal{B}_W$ as linear combinations of the basis vectors $\mathcal{B}_V$ of Example 1.

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 3 \\ 5 & 0 & 2 \\ -1 & 3 & -3 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \\ 3 \end{pmatrix} = \mathbf{V}\mathbf{s}_1$$

$$\begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 3 \\ 5 & 0 & 2 \\ -1 & 3 & -3 \end{pmatrix} \begin{pmatrix} -30/7 \\ 58/7 \\ 68/7 \end{pmatrix} = \mathbf{V}\mathbf{s}_2.$$

$$\begin{pmatrix} 8 \\ 1 \\ -10 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 3 \\ 5 & 0 & 2 \\ -1 & 3 & -3 \end{pmatrix} \begin{pmatrix} -5/7 \\ -9/7 \\ 16/7 \end{pmatrix} = \mathbf{V}\mathbf{s}_3.$$

It then follows that $\mathbf{W} = \mathbf{VS}$ is

$$\begin{pmatrix} 1 & 4 & 8 \\ 1 & -2 & 1 \\ 1 & 0 & -10 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 3 \\ 5 & 0 & 2 \\ -1 & 3 & -3 \end{pmatrix} \begin{pmatrix} -1 & -30/7 & -5/7 \\ 3 & 58/7 & -9/7 \\ 3 & 68/7 & 16/7 \end{pmatrix}.$$

The matrix $\mathbf{S}$ is called the <u>transition matrix</u> from the basis $\mathcal{B}_W$ to the basis $\mathcal{B}_V$.

Example 6. Consider the bases $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \mathbf{w}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Find the transition matrix $\mathbf{S}$ from basis $\mathcal{B}_W$ to the basis $\mathcal{B}_V$.

Solution. According to (37), we must find the coordinates of the vectors in $\mathcal{B}_W$ with respect to the basis $\mathcal{B}_V$. We have

$$\mathbf{w}_1 = s_{11}\mathbf{v}_1 + s_{21}\mathbf{v}_2 = 1 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2$$

and

$$\mathbf{w}_2 = s_{12}\mathbf{v}_1 + s_{22}\mathbf{v}_2 = 2 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2.$$

The transition matrix $\mathbf{S}$ is

$$\mathbf{S} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

Suppose again that both $\mathbf{V}$ and $\mathbf{W}$ are $n \times k$ matrices, not necessarily square matrices. Any vector $\mathbf{v}$ admits the following representations relative to the two bases $\mathcal{B}_V$ and $\mathcal{B}_W$ respectively

$$\mathbf{v} = \mathbf{V}\mathbf{x}_V \qquad \text{and} \qquad \mathbf{v} = \mathbf{W}\mathbf{x}_W.$$

Formula (38) establishes a relation between the two bases in such a way as to allows us to write the second equality above as

$$\mathbf{v} = \mathbf{W}\mathbf{x}_W = \mathbf{V}\mathbf{S}\mathbf{x}_W. \tag{39}$$

Since $\mathbf{v}$ is given by either

$$\mathbf{v} = \mathbf{V}\mathbf{x}_V \qquad \text{or} \qquad \mathbf{v} = \mathbf{V}(\mathbf{S}\mathbf{x}_W),$$

we conclude that $\mathbf{S}\mathbf{x}_W$ is also the coordinate vector of $\mathbf{v}$ relative to the basis $\mathcal{B}_V$. But since the coordinate vector in a specific basis is unique we must have

$$\mathbf{x}_V = \mathbf{S}\mathbf{x}_W. \tag{40}$$

This is the connection between the coordinate vectors in the two bases. Since we have already proven that $\mathbf{S}$ is square and invertible, it follows from (40) that

$$\mathbf{x}_W = \mathbf{S}^{-1}\mathbf{x}_V. \tag{41}$$

Example 7. Verify formula (41) for the bases $\mathcal{B}_V$ and $\mathcal{B}_W$ of the previous examples.

Solution. The matrix $\mathbf{S}$ is given in example 5. Thus

$$\begin{pmatrix} 1 \\ -1/5 \\ -2/5 \end{pmatrix} = \begin{pmatrix} -1 & -30/7 & -5/7 \\ 3 & 58/7 & -9/7 \\ 3 & 68/7 & 16/7 \end{pmatrix}^{-1} \begin{pmatrix} 1/7 \\ 13/7 \\ 1/7 \end{pmatrix}.$$

1) Consider the bases $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2\}$ for R^2, where

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

a) Use (32) to find the coordinate vector of $\mathbf{v} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$ relative to $\mathcal{B}_V$.

b) Use (32) to find the coordinate vector of $\mathbf{v}$ relative to $\mathcal{B}_W$.

c) Find the transition matrix $\mathbf{S}$ from basis $\mathcal{B}_W$ to the basis $\mathcal{B}_V$.

d) Find the transition matrix **S'** from basis $\mathcal{B}_V$ to the basis $\mathcal{B}_W$.

e) Verify formulas (63) and (64).

2) Consider a first basis $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for a subspace of R^4, where

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix},$$

and a second basis $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ for the same subspace, where

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 4 \end{pmatrix} \qquad \mathbf{w}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -2 \end{pmatrix}.$$

a) Write down formula (38) for the given bases.

b) Find the coordinates of the vectors $\mathbf{b} = \begin{pmatrix} 3 \\ 7 \\ 3 \\ 10 \end{pmatrix}$ relative to the basis $\mathcal{B}_V$.

c) Use (64) to find the coordinates of $\mathbf{b}$ relative $\mathcal{B}_W$.

8.3 Linear Transformations

Suppose now that the vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{V}$ are related through the equation

$$\mathbf{y} = T(\mathbf{x}). \tag{42}$$

We say that T is a function that carries the vector $\mathbf{x}$ into the vector $\mathbf{y}$ in the same sense that $y = f(x) = x^2$ is a function that assigns a value $f(x)$ to each independent variable x.
The difference is that in (42) the independent variable is a vector $\mathbf{x}$ in a k-dimensional subspace $\mathcal{V}$ and the dependent variable is another vector $\mathbf{y}$ in the same subspace. In this case we write $T : \mathcal{V} \to \mathcal{V}$. We will be interest in a special class of such functions called **linear maps or linear transformations**. For these functions, the following two conditions hold for all vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{V}$, and any scalar k:

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= T(\mathbf{x}) + T(\mathbf{y}) \\ T(k \cdot \mathbf{x}) &= k \cdot T(\mathbf{x}). \end{aligned} \tag{43}$$

Example 8. Define $T : R^2 \to R^2$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ x_1 \end{pmatrix}.$$

T is linear transformation because it satisfies the two conditions (43) above. On the other hand, the map $T : R^2 \to R^2$ defined by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ x_1 x_2 \end{pmatrix}$$

is not linear, since it fails to satisfy the two conditions (43).

Example 9. Define $T : R^2 \to R^2$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}.$$

T is called the **reflection about the x-axis**.

Example 10. Define $T : R^2 \to R^2$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 \\ x_2 \end{pmatrix}.$$

T is called the **reflection about the y-axis**.

Example 11. Define $T_\theta : R^2 \to R^2$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}.$$

T_θ is called the **reflection about the line** $x_1 = x_2$.

Example 12. Define $T : R^2 \to R^2$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}.$$

T is called the **projection on the x-axis**.

Example 13. For any angle θ, define $T_\theta : R^2 \to R^2$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \cos\theta - x_2 \sin\theta \\ x_1 \sin\theta + x_2 \cos\theta \end{pmatrix}.$$

T_θ is called the **rotation by** θ.

8.4 The matrix of a linear transformation

Suppose now that $\mathbf{x}_V$ and $\mathbf{y}_V$ are the coordinate vectors of the vectors $\mathbf{x}$ and $\mathbf{y} = T(\mathbf{x})$ relative to the basis $\mathcal{B}_V$. We first want to investigate how the two coordinates of these vectors are related relative to the basis $\mathcal{B}_V$. Next, we want to determine how the transformation is represented in different bases. Relative to the basis $\mathcal{B}_V$ the vector $\mathbf{v}$ is given by (30):

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_k\mathbf{v}_k \tag{44}$$

The function T applied to $\mathbf{x}$ is

$$\begin{aligned} T(\mathbf{v}) &= T(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_k\mathbf{v}_k) \\ &= T(x_1\mathbf{v}_1) + T(x_2\mathbf{v}_2) + ... + T(x_k\mathbf{v}_k) \\ &= x_1T(\mathbf{v}_1) + x_2T(\mathbf{v}_2) + ... + x_kT(\mathbf{v}_k), \end{aligned} \tag{45}$$

where we have used the two linearity conditions (43).
In matrix notation (45) is

$$T(\mathbf{v}) = [T(\mathbf{v}_1) \;\; T(\mathbf{v}_2) \;\; ... \; T(\mathbf{v}_k)] \; \mathbf{x}_V. \tag{46}$$

This result shows that effect of the linear transformation on $\mathbf{v}$ is completely determined by knowing how the linear transformation transforms each of the basis vectors. Since each $T(\mathbf{v}_j)$ is a vector in $\mathcal{V}$, each of them can be written as a linear combination of the basis vectors. Thus, let the effect of the linear transformation on the basis vectors be represented by

$$\begin{aligned} T(\mathbf{v}_1) &= a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + ... + a_{k1}\mathbf{v}_k \\ T(\mathbf{v}_2) &= a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + ... + a_{k2}\mathbf{v}_k \\ ... & \qquad ... \qquad\qquad ... \\ T(\mathbf{v}_k) &= a_{1k}\mathbf{v}_1 + a_{2k}\mathbf{v}_2 + ... + a_{kk}\mathbf{v}_k, \end{aligned} \tag{47}$$

in close analogy to (37).
Since each of the $T(\mathbf{v}_j)'s$ can be put in matrix notation analogous to that

used in (38), we also have

$$T(\mathbf{v}_1) \quad = \quad \begin{pmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_k \\ | & | & & | \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{k1} \end{pmatrix} = \mathbf{V} \cdot \mathbf{a}_1$$

$$\dots \qquad \dots \qquad \dots \qquad \dots$$

$$T(\mathbf{v}_k) \quad = \quad \begin{pmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_k \\ | & | & & | \end{pmatrix} \begin{pmatrix} a_{1k} \\ a_{2k} \\ \cdot \\ \cdot \\ \cdot \\ a_{kk} \end{pmatrix} = \mathbf{V} \cdot \mathbf{a}_k.$$

It is now clear that

$$\begin{aligned} [T(\mathbf{v}_1) \;\; T(\mathbf{v}_2) \;\; \dots \;\; T(\mathbf{v}_k)] &= [\mathbf{V}\mathbf{a}_1 \;\; \mathbf{V}\mathbf{a}_2 \;\; \dots \;\; \mathbf{V}\mathbf{a}_k] \\ &= \mathbf{V} \; [\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \dots \;\; \mathbf{a}_k] \\ &= \mathbf{V}\mathbf{A}_V \end{aligned} \tag{48}$$

for some $k \times k$ matrix $\mathbf{A}_V$. This matrix is called the *matrix of the transformation T with respect to the basis $\mathcal{B}_V$.* Note that $\mathbf{A}_V$ is just the transpose of the matrix of the coefficients in (47).
Substituting (48) into (46) the transformed vector $\mathbf{y} = T(\mathbf{v})$ can be represented in the following way relative to the basis $\mathcal{B}_V$.

$$\mathbf{y} = T(\mathbf{v}) = \mathbf{V}\mathbf{A}_V\mathbf{x}_V.$$

By setting $\mathbf{y} = \mathbf{V}\mathbf{y}_V$ relative to this same basis, we conclude that

$$\mathbf{y}_V = \mathbf{A}_V\mathbf{x}_V. \tag{49}$$

Example 14. Let $T : R^2 \to R^2$ be a linear transformation defined by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 - 3x_2 \\ 3x_1 + 5x_2 \end{pmatrix}.$$

a) Find the matrix of the transformation with respect to the basis $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

b) Find the matrix of the transformation with respect to the basis $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Solution a) According to (47) we have to determine the effect of the T on the basis vectors of $\mathcal{B}_V$. We have

$$T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and

$$T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 5 \end{pmatrix} = -3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus,

$$\mathbf{A}_V = \begin{pmatrix} 2 & -3 \\ 3 & 5 \end{pmatrix}.$$

b) Let us write down the set of equation (47) for the given T. We have

$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 8 \end{pmatrix} = \frac{7}{2} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{9}{2} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and

$$T\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \end{pmatrix} = \frac{3}{2} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{7}{2} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Thus,

$$\mathbf{A}_W = \begin{pmatrix} 7/2 & 3/2 \\ -9/2 & -7/2 \end{pmatrix}.$$

Example 15. Let $T : R^2 \to R^2$ be a linear transformation defined by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 \end{pmatrix}.$$

a) Find the matrix of the transformation with respect to the basis $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

b) Find the matrix of the transformation with respect to the basis $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Solution a) According to (47) we have

$$T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and

$$T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus,

$$\mathbf{A}_V = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

b) Let us write down the set (47) for this basis.

$$T\begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} - 3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = -1 \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Thus,

$$\mathbf{A}_W = \begin{pmatrix} 0 & -1 \\ -3 & 2 \end{pmatrix}.$$

3) Let $T : R^2 \to R^2$ be a linear transformation defined by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + 2x_2 \\ -5x_1 + 4x_2 \end{pmatrix}.$$

a) Find the matrix of the transformation with respect to the basis $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

b) Find the matrix of the transformation with respect to the basis $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

4) Let $T : R^2 \to R^2$ be a linear transformation defined by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 \end{pmatrix}.$$

a) Find the matrix of the transformation with respect to the basis $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

b) Find the matrix of the transformation with respect to the basis $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

8.5 Similarity Transformations

We now tackle the question of how the matrix of a linear transformation relative to a certain basis is related to the matrix of the same linear transformation relative to a second basis for the same subspace. From equation (49) we know that

$$\mathbf{y}_V = \mathbf{A}_V \mathbf{x}_V \tag{50}$$

holds with respect to $\mathcal{V}$.
It is clear that in a second basis $\mathcal{B}_W$ we will have, in a simliar fashion,

$$\mathbf{y}_W = \mathbf{A}_W \mathbf{x}_W. \tag{51}$$

Note now that we know from (63) that the coordinate vectors relative to these two bases are related through the equations

$$\mathbf{y}_V = \mathbf{S}\mathbf{y}_W \qquad \text{and} \qquad \mathbf{x}_V = \mathbf{S}\mathbf{x}_W,$$

where $\mathbf{S}$ is the matrix appearing in (37). Substituting these expression in (50) yields

$$\mathbf{S}\mathbf{y}_W = \mathbf{A}\mathbf{S}\mathbf{x}_W,$$

and hence

$$\mathbf{y}_W = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}\mathbf{x}_W. \tag{52}$$

since $\mathbf{S}$ is an invertible matrix.
We then conclude from (52) and (51) that

$$\mathbf{A}_W = \mathbf{S}^{-1}\mathbf{A}_V\mathbf{S}.$$

In general if there exists a non-singular matrix $\mathbf{S}$ such that $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{B}$, the matrix $\mathbf{B}$ is said to be **similar** to $\mathbf{A}$, and we say that $\mathbf{B}$ is obtained from $\mathbf{A}$ by a **similarity transformation**.

Example 16. Let $\mathcal{B}_V$ and $\mathcal{B}_W$ be the two bases for R^2 given in example 15. Suppose that the coordinate vectors of $\mathbf{x}$ and $\mathbf{y}$ are related according to the equation

$$\mathbf{y}_V = \mathbf{A}\mathbf{x}_V = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \mathbf{x}_V$$

relative to the basis $\mathcal{B}_V$. To find the coordinate vectors of $\mathbf{x}$ and $\mathbf{y}$ relative to $\mathcal{B}_W$ we have to obtain the matrix $\mathbf{S}$. This matrix is given by (37) or (38) as a connection between the vectors of the two bases. Let us write the each vector of $\mathcal{B}_W$ as a linear combination of the vectors of $\mathbf{B}_V$. We have

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus,

$$\mathbf{S} = \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix}.$$

Then,

$$\mathbf{y}_W = \mathbf{B}\mathbf{x}_W = \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix} \mathbf{x}_W = \begin{pmatrix} 0 & -1 \\ -3 & 2 \end{pmatrix} \mathbf{x}_W,$$

which is the matrix of the linear transformation relative to $\mathcal{B}_W$ that we have obtained by direct computation in example 15.

Example 17. The matrices

$$\mathbf{A} = \begin{pmatrix} 6 & 2 \\ -2 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 8 & 6 \\ -3 & -1 \end{pmatrix}$$

are similar since there is a matrix $\mathbf{S} = \begin{pmatrix} 3 & 2 \\ 0 & 2 \end{pmatrix}$ such that

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \begin{pmatrix} 1/3 & -1/3 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 6 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 6 \\ -3 & -1 \end{pmatrix}.$$

5) Let $\mathcal{B}_V$ and $\mathcal{B}_W$ be the two bases for R^2 given in example 6. Suppose that the matrix of a linear transformation relative to the basis $\mathcal{B}_V$ is

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}.$$

Find the matrix of the representation relative to the basis $\mathcal{B}_W$.

6) Let $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

be a basis for R^2. The matrix of a linear transformation T with respect to $\mathcal{B}_V$ is given by

$$\begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}.$$

Let $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{w}_1 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

be another basis for R^2. What is the matrix of the linear transformation T with respect to the basis $\mathcal{B}_W$?

7) Prove that if $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, then $det(\mathbf{B}) = det(\mathbf{A})$.

8) Prove that similar matrices have the same characteristic equation and the same eigenvalues.

The usefulness of a similarity transformation lies in the fact that it can be applied to a matrix $\mathbf{A}$ to obtain a similar matrix $\mathbf{B}$ that has a simpler structure than $\mathbf{A}$. From now on let us take $\mathbf{A}$ to be an $n \times n$ matrix.
There are basically three special types of matrices that can be obtained by a similarity transformation applied to $\mathbf{A}$. For a suitable $\mathbf{S}$, the resulting matrix $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ can be reduced to a diagonal matrix, an upper-triangular matrix or a block-diagonal matrix.

In this assignment we will be interested in similarity transformations that produce diagonal matrices. In our next assignment we are going to deal with transformations that produce upper-triangular matrices and, if time allows, we will discuss block-diagonal matrices.

8.6 Reduction to diagonal form.

Let us start by investigating under which conditions there exists a non-singular matrix $\mathbf{S}$ such that the similarity transformation $\mathbf{S}^{-1}\mathbf{AS}$ produces diagonal matrix $\mathbf{B}$. In this case we say that the matrix $\mathbf{A}$ has been reduced to diagonal form, or that $\mathbf{A}$ is **diagonalizable**. The matrix $\mathbf{S}$ is said to diagonalize $\mathbf{A}$.
The reduction to diagonal form is an important procedure whenever an equation of the form

$$\mathbf{y}_V = \mathbf{A}\mathbf{x}_V,$$

or, as we will see, variations of it, is deemed too complicated with respect to the basis $\mathcal{B}_V$. Then it might be more convenient to pass from the basis $\mathcal{B}_V$ to a second basis $\mathcal{B}_W$ in such a way that in this second basis the corresponding equation

$$\mathbf{y}_W = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}\mathbf{x}_W = \mathbf{\Lambda}\mathbf{x}_W \tag{53}$$

assumes a particularly simple form since $\mathbf{\Lambda} = \mathbf{S}^{-1}\mathbf{AS}$ is diagonal.
If we can find such an $\mathbf{S}$ and are able to solve whatever problem related to equation (49) using the basis $\mathcal{B}_W$, then we can always go back to the original basis $\mathcal{B}_V$ since the two basis are connected via formula (38).
The key point that makes the reduction to diagonal form possible is that the matrix $\mathbf{A}$ has n linearly independent eigenvectors. Any square matrix of size n that has n linearly independent eigenvectors can be reduced to diagonal matrix by a similarity transformation.
The reduction of a matrix $\mathbf{A}$ to diagonal form is carried out by setting

$$\mathbf{S} = \begin{pmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ | & | & & | \end{pmatrix}$$

where $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ are the eigenvectors of $\mathbf{A}$ corresponding to the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ respectively. Then note that

$$\begin{aligned} \mathbf{AS} &= \mathbf{A}\,[\mathbf{x}_1 \;\; \mathbf{x}_2 \;\; ... \;\; \mathbf{x}_n] \\ &= [\mathbf{A}\mathbf{x}_1 \;\; \mathbf{A}\mathbf{x}_2 \;\; ... \;\; \mathbf{A}\mathbf{x}_n] \\ &= [\lambda_1\mathbf{x}_1 \quad \lambda_2\mathbf{x}_2 \quad ... \quad \lambda_n\mathbf{x}_n] \\ &= \mathbf{S}\mathbf{\Lambda}, \end{aligned} \tag{54}$$

where $\mathbf{\Lambda}$ is a diagonal matrix having the eigenvalues of $\mathbf{A}$ as diagonal elements.

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ . & . & . & . \\ 0 & 0 & ... & \lambda_n \end{pmatrix}.$$

It follows from (54) then that $\mathbf{\Lambda} = \mathbf{S}^{-1}\mathbf{AS}$.
Note that it is not necessary that the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ be distinct.

Example 18. Let $\mathcal{V}$ be a two-dimensional vector space, and let $\mathbf{A}$ be the matrix

$$\mathbf{A} = \begin{pmatrix} -2 & 4 \\ 1 & 1 \end{pmatrix}.$$

The characteristic polynomial of $\mathbf{A}$ is

$$p(\lambda) = det(\mathbf{A} - \lambda\mathbf{I}) = \lambda^2 + \lambda - 6,$$

and thus the eigenvalues of $\mathbf{A}$ are $\lambda_1 = 2$ and $\lambda_2 = -3$.

To find the eigenvector corresponding to $\lambda_1 = 2$ we solve the equation $(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} = 0$, or

$$\begin{pmatrix} -4 & 4 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{0}.$$

From this equation we obtain $x_1 = x_2$. Thus, any vector of the form

$$\begin{pmatrix} x_1 \\ x_1 \end{pmatrix} \qquad x_1 \neq 0$$

is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalues λ_1. For convenience, let us choose $x_1 = 1$. Then

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

is an eigenvector.
In a similar fashion, an eigenvector $\mathbf{x}_2$ corresponding to the eigenvalue $\lambda_2 = -3$ is

$$\mathbf{x}_2 = \begin{pmatrix} 4 \\ -1 \end{pmatrix}.$$

Now, construct the matrix

$$\mathbf{S} = [\mathbf{x}_1 \;\; \mathbf{x}_2] = \begin{pmatrix} 1 & 4 \\ 1 & -1 \end{pmatrix} \qquad \text{with} \qquad \mathbf{S}^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 4 \\ 1 & -1 \end{pmatrix}.$$

Then, the similarity transformation $\mathbf{\Lambda} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ produces a diagonal matrix

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \frac{1}{5}\begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -2 & 4 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

having the eigenvalues of $\mathbf{A}$ as diagonal elements.

If the matrix $\mathbf{A}$ above is given with respect to a basis $\mathcal{B}_V = \{\mathbf{v}_1 \;\; \mathbf{v}_2\}$, the diagonal matrix $\mathbf{\Lambda} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is given with respect to a second basis $\mathcal{B}_W = \{\mathbf{w}_1 \;\; \mathbf{w}_2\}$. The connection between the two bases is given (38), from which we get

$$\mathbf{w}_1 = \mathbf{v}_1 + \mathbf{v}_2 \qquad \text{and} \qquad \mathbf{w}_2 = 4\mathbf{v}_1 - \mathbf{v}_2.$$

In the above example, the first column of the matrix $\mathbf{S}$ is a particular eigenvector corresponding to the eigenvalue $\lambda_1 = 2$, and the second column of $\mathbf{S}$ is a particular eigenvector corresponding to the eigenvector $\lambda_2 = -3$.
What would happen if we chose different eigenvectors corresponding to the eigenvalues 2 and -3? The answer is ... : Nothing.

Any matrix whose columns are linearly independent eigenvectors will diagonalize $\mathbf{A}$. For example,

$$\mathbf{x}_1 = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{pmatrix} 20 \\ -5 \end{pmatrix}$$

are also eigenvector of $\mathbf{A}$ associated with 2 and -3 respectively. If we construct the following new $\mathbf{S}$

$$\mathbf{S} = \begin{pmatrix} 3 & 20 \\ 3 & -5 \end{pmatrix} \quad \text{with} \quad \mathbf{S}^{-1} = \begin{pmatrix} 1/15 & 4/15 \\ 1/25 & -1/25 \end{pmatrix},$$

then

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \begin{pmatrix} 1/15 & 4/15 \\ 1/25 & -1/25 \end{pmatrix} \begin{pmatrix} -2 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 20 \\ 3 & -5 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$$

as well.

What would happen if we wrote the eigenvectors in different order as columns of a new matrix $\mathbf{S}$? The answer is that the new $\mathbf{S}$ would still diagonalize $\mathbf{A}$, but the eigenvalues appearing on the main diagonal of the result matrix would appear in the order corresponding to the eigenvectors in the columns of $\mathbf{S}$. For example, let

$$\mathbf{S} = \begin{pmatrix} 20 & 3 \\ -5 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{S}^{-1} = \begin{pmatrix} 1/25 & -1/25 \\ 1/15 & 4/25 \end{pmatrix},$$

then

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \begin{pmatrix} 1/25 & -1/15 \\ 1/15 & 4/25 \end{pmatrix} \begin{pmatrix} -2 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 20 & 3 \\ -5 & 3 \end{pmatrix} = \begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix}.$$

21) Find an orthogonal matrix $\mathbf{S}$ such that $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is diagonal, where $\mathbf{A}$ is the symmetric matrix

$$\mathbf{A} = \begin{pmatrix} 7 & -16 & -8 \\ -16 & 7 & 8 \\ -8 & 8 & -5 \end{pmatrix}.$$

22) Compute $\mathbf{A}^8 - 3\mathbf{A}^5 + \mathbf{A}^4 + \mathbf{A}^2 - 4\mathbf{I}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

23) When will the matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

have equal eigenvalues?

24) Show that the characteristic equation of the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{pmatrix}$$

is $\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$.

25) Find a non-singular matrix $\mathbf{S}$ that diagonalizes the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 4 \\ 0 & 3 \end{pmatrix}.$$

26) Find a non-singular matrix $\mathbf{S}$ that diagonalizes the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 4 \\ 3 & -2 \end{pmatrix}.$$

27) Explain why the matrices

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

fail to be diagonalizable.

28) Reduce the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{pmatrix}$$

to diagonal form.

29) Find an orthogonal reduction (that is, $\mathbf{S}$ is orthogonal) to diagonal form for the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{pmatrix}.$$

We can illustrate the usefulness of the reduction to diagonal form by showing how it can be used to find $\mathbf{A}^k$ for arbitrary positive integral k.
If a matrix $\mathbf{S}$ exists such that $\mathbf{A} = \mathbf{S\Lambda S}^{-1}$, where Λ is diagonal, then

$$\mathbf{A}^k = (\mathbf{S\Lambda S}^{-1}) \cdot (\mathbf{S\Lambda S}^{-1}) \cdot (\mathbf{S\Lambda S}^{-1}) \ldots (\mathbf{S\Lambda S}^{-1}) \cdot (\mathbf{S\Lambda S}^{-1})$$

where the product $(\mathbf{S\Lambda S}^{-1})$ is repeated k times. But

$$\begin{aligned} \mathbf{A}^k &= (\mathbf{S\Lambda S}^{-1}) \cdot (\mathbf{S\Lambda S}^{-1}) \cdot (\mathbf{S\Lambda S}^{-1}) \ldots (\mathbf{S\Lambda S}^{-1}) \cdot (\mathbf{S\Lambda S}^{-1}) \\ &= \mathbf{S\Lambda} \cdot (\mathbf{S}^{-1}\mathbf{S}) \cdot \mathbf{\Lambda} \cdot (\mathbf{S}^{-1}\mathbf{S}) \ldots (\mathbf{S}^{-1}\mathbf{S}) \cdot \mathbf{\Lambda S}^{-1} \\ &= \mathbf{S\Lambda}^k\mathbf{S}^{-1}. \end{aligned}$$

Since $\mathbf{\Lambda}$ is a diagonal matrix with diagonal elements λ_j, $\mathbf{\Lambda}^k$ is a diagonal matrix with diagonal element λ_j^k.

30) Find $\mathbf{A}^k$ for any integer k for the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

31) Find $\mathbf{A}^{25}$, where

$$\mathbf{A} = \begin{pmatrix} -1 & 2 \\ 4 & 1 \end{pmatrix}.$$

9 Schur's Theorem.

If an $n \times n$ matrix $\mathbf{A}$ has n distinct eigenvalues it must have n linearly independent eigenvectors. This matrix can be reduced to diagonal form by a similarity transformation. If $\mathbf{A}$ has **repeated eigenvalues**, it may or may not have n linearly independent eigenvectors. If it does, then it can also be reduced to diagonal form by a similarity transformation. If it does not have n linearly independent eigenvectors, then it cannot be reduced to diagonal form.

Example 1. The matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & -2 \\ 0 & 4 & -2 \\ 0 & 3 & -1 \end{pmatrix}$$

has characteristic equation $(1-\lambda)^2(2-\lambda) = 0$. The vectors

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

are two linearly independent eigenvectors corresponding to $\lambda_1 = 1$. Corresponding to $\lambda_2 = 2$ there is the eigenvector

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

By letting

$$\mathbf{S} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{pmatrix}$$

the matrix $\mathbf{A}$ can be diagonalized by the transformation $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$.

Example 2. The matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & -2 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

has characteristic equation $(1-\lambda)(2-\lambda)^2 = 0$. The eigenvector corresponding to $\lambda_1 = 1$ is

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

There is only one eigenvector corresponding to $\lambda_2 = 2$:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Since there are only two linearly independent eigenvectors we are not able to diagonalize $\mathbf{A}$.

The characteristic polynomial of an $n \times n$ $\mathbf{A}$ can always be factored in the following way

$$det(\mathbf{A} - \lambda\mathbf{I}) = (\lambda - \lambda_1)^{a_1} \cdot (\lambda - \lambda_2)^{a_2} \ . \ . \ . \ (\lambda - \lambda_p)^{a_p} \ ,$$

where λ_j , for $j = 1, 2, ..., p$, are the distinct eigenvalues of $\mathbf{A}$.

The eigenvalue λ_1 is repeated a_1 times as a root of the characteristic polynomial; the eigenvalue λ_2 is repeated a_2 times, ..., and the eigenvalue λ_p is repeated a_p times as a root of the characteristic polynomial.

Each number a_j is called the **algebraic multiplicity** of the eigenvalue λ_j. The a_j are positive integers and $a_1 + a_2 + \ . \ . \ . \ + a_p = n$.

Let $g(\lambda_j)$ be the number of eigenvectors corresponding to an eigenvalue λ_j that is repeated a_j times. We call $g(\lambda_j)$ the **geometric multiplicity** of that eigenvalue. It is always true that $g(\lambda_j) \leq a_j$.

In case the matrix $\mathbf{A}$ has no repeated eigenvalues, then each $a_j = 1$ for all j and $p = n$. Then $\mathbf{A}$ has n linearly independent eigenvectors. If $\mathbf{A}$ has repeated eigenvalues, then the algebraic multiplicity of one or more of the eigenvalues is greater than one. In this case it is not always possible to diagonalize the matrix because it may not have n linearly independent eigenvectors. However, if $g(\lambda_j) = a_j$ for all j, then diagonalization is always possible.

In this assignment we are going to study $n \times n$ matrices that fail to have n linearly independent eigenvectors, and thus cannot be reduced to diagonal form through a similarity transformation. These matrices admit, however, a reduction to upper-triangular form via a similarity transformation.

9.1 Reduction to upper-triangular form.

We will show that any $n \times n$ matrix $\mathbf{A}$ is similar to an upper-triangular matrix. The reduction of a matrix $\mathbf{A}$ to upper-triangular form is carried out by setting

$$\mathbf{S}_1 = \begin{pmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{y}_2 & \dots & \mathbf{y}_n \\ | & | & & | \end{pmatrix}$$

where $\mathbf{x}_1$ is a normalized eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ_1, and the normalized vectors $\mathbf{y}_2, ..., \mathbf{y}_n$ are totally arbitrary vectors except for the condition that $\mathbf{x}_1, \mathbf{y}_2, ..., \mathbf{y}_n$ form an orthogonal basis.

Then note that

$$\begin{aligned} \mathbf{A}\mathbf{S}_1 &= \mathbf{A}\,[\mathbf{x}_1 \;\; \mathbf{y}_2 \;\; \dots \;\; \mathbf{y}_n] \\ &= [\mathbf{A}\mathbf{x}_1 \;\; \mathbf{A}\mathbf{y}_2 \;\; \dots \;\; \mathbf{A}\mathbf{y}_n] \\ &= [\lambda_1 \mathbf{x}_1 \quad \mathbf{A}\mathbf{y}_2 \quad \dots \quad \mathbf{A}\mathbf{y}_n] \end{aligned}$$

where we have used that $\mathbf{A}\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$.

Since the columns of $\mathbf{S}_1$ are linearly independent vector, $\mathbf{S}_1$ is an orthogonal matrix. Therefore $\mathbf{S}_1^{-1} = \mathbf{S}_1^T$. It follows then that

$$
\begin{aligned}
\mathbf{S}^{-1}\mathbf{A}\mathbf{S}_1 &= \mathbf{S}_1^T\mathbf{A}\mathbf{S}_1 \\
&= \begin{pmatrix} -- & \mathbf{x}_1 & -- \\ -- & \mathbf{y}_2 & -- \\ . & . & . \\ -- & \mathbf{y}_n & -- \end{pmatrix} \begin{pmatrix} | & | & & | \\ \lambda_1\mathbf{x}_1 & \mathbf{y}_2 & ... & \mathbf{y}_n \\ | & | & & | \end{pmatrix} \\
&= \begin{pmatrix} \langle \mathbf{x}_1, \lambda_1\mathbf{x}_1\rangle & \langle \mathbf{x}_1, \mathbf{A}\mathbf{y}_2\rangle & ... & \langle \mathbf{x}_1, \mathbf{A}\mathbf{y}_n\rangle \\ \langle \mathbf{y}_2, \lambda_1\mathbf{x}_1\rangle & \langle \mathbf{y}_2, \mathbf{A}\mathbf{y}_2\rangle & ... & \langle \mathbf{y}_2, \mathbf{A}\mathbf{y}_n\rangle \\ . & . & ... & . \\ \langle \mathbf{y}_n, \lambda_1\mathbf{x}_1\rangle & \langle \mathbf{y}_n, \mathbf{A}\mathbf{y}_2\rangle & ... & \langle \mathbf{y}_n, \mathbf{A}\mathbf{y}_n\rangle \end{pmatrix}. \qquad (55)
\end{aligned}
$$

We have assumed that $\mathbf{x}_1$ is a normalized vector and that it is orthogonal to $\mathbf{y}_2, ..., \mathbf{y}_n$. Thus, the first column of the matrix (55) has only one nonzero entry:

$$
\mathbf{S}_1^{-1}\mathbf{A}\mathbf{S}_1 = \begin{pmatrix} \lambda_1 & \langle \mathbf{x}_1, \mathbf{A}\mathbf{y}_2\rangle & ... & \langle \mathbf{x}_1, \mathbf{A}\mathbf{y}_n\rangle \\ 0 & \langle \mathbf{y}_2, \mathbf{A}\mathbf{y}_2\rangle & ... & \langle \mathbf{y}_2, \mathbf{A}\mathbf{y}_n\rangle \\ . & . & ... & . \\ 0 & \langle \mathbf{y}_n, \mathbf{A}\mathbf{y}_2\rangle & ... & \langle \mathbf{y}_n, \mathbf{A}\mathbf{y}_n\rangle \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & \mathbf{B} \end{pmatrix}, \qquad (56)
$$

where $*$ stands for unspecified entries and $\mathbf{B}$ is the $(n-1)\times(n-1)$ matrix

$$
\mathbf{B} = \begin{pmatrix} \langle \mathbf{y}_2, \mathbf{A}\mathbf{y}_2\rangle & ... & \langle \mathbf{y}_2, \mathbf{A}\mathbf{y}_n\rangle \\ . & ... & . \\ \langle \mathbf{y}_n, \mathbf{A}\mathbf{y}_2\rangle & ... & \langle \mathbf{y}_n, \mathbf{A}\mathbf{y}_n\rangle \end{pmatrix} \qquad (57)
$$

Form (56) is the first step in the process to reduce $\mathbf{A}$ to upper-triangular form via a similarity transformation. At the end of this first process we have produce zeros below the first entry of the first column of $\mathbf{A}$.

If $\mathbf{B}$ is already in upper-triangular form the process stops here and (56) is the similarity transformation that reduces $\mathbf{A}$ to upper-triangular form.

If $\mathbf{B}$ is not in upper-triangular form, we start the second step of the reduction process. The second step runs as follows. Repeat the first step described above for the matrix $\mathbf{B}$ in (57). Choose a normalized eigenvector $\mathbf{x}_2$ of $\mathbf{B}$ corresponding to an eigenvalue λ_2. Construct the matrix

$$\mathbf{S} = \begin{pmatrix} | & | & & | \\ \mathbf{x}_2 & \mathbf{z}_3 & \dots & \mathbf{z}_n \\ | & | & & | \end{pmatrix}$$

where $\mathbf{z}_3, ..., \mathbf{z}_n$ are arbitrary vectors, except for the condition the they are normalized and $\mathbf{x}_2, \mathbf{z}_3, ..., \mathbf{z}_n$ form an orthogonal basis. This condition ensures that $\mathbf{S}$ is an orthogonal matrix and thus $\mathbf{S}^{-1} = \mathbf{S}^T$. At the end of the second step we arrive at an expression for $\mathbf{S}^{-1}\mathbf{B}\mathbf{S}$ of the same type as (56).

$$\mathbf{S}^{-1}\mathbf{B}\mathbf{S} = \begin{pmatrix} \lambda_2 & * \\ \mathbf{0} & \mathbf{C} \end{pmatrix}$$

where $\mathbf{C}$ is an $(n-2) \times (n-2)$ matrix.
Next construct the matrix

$$\mathbf{S}_2 = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{pmatrix}$$

and note that $\mathbf{S}_2^{-1}(\mathbf{S}_1^{-1}\mathbf{A}\mathbf{S}_1)\mathbf{S}_2 = (\mathbf{S}_1\mathbf{S}_2)^{-1}\mathbf{A}(\mathbf{S}_1\mathbf{S}_2)$ is a similarity transformation for $\mathbf{T} = \mathbf{S}_1\mathbf{S}_2$. Thus

$$\begin{aligned} \mathbf{T}^{-1}\mathbf{A}\mathbf{T} &= \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & \mathbf{S}^{-1}\mathbf{B}\mathbf{S} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & * & * \\ \mathbf{0} & \lambda_2 & * \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \end{pmatrix}. \end{aligned} \qquad (58)$$

If the $(n-2) \times (n-2)$ matrix $\mathbf{C}$ is in upper-triangular form the process stops here. Otherwise we repeat the steps above for the matrix $\mathbf{C}$. It is easy to see now that this succession of transformations is intended to clear the column entries below the eigenvalue λ_1 of $\mathbf{A}$, then the eigenvalue λ_2 of $\mathbf{B}$, then the eigenvalue λ_3 of $\mathbf{C}$ and so forth, thereby obtaining an upper triangular matrix at the very end of the process.

The important point to observe here is that the eigenvalues of the matrices $\mathbf{B}$, $\mathbf{C}$ and all the subsequent matrices that are generated in the intermediate steps of the triangularization process are all eigenvalues of $\mathbf{A}$. Thus, in (58) above, λ_1 and λ_2 are indeed two eigenvalues of $\mathbf{A}$! It is pretty easy to see this. Recall from (56) that

$$\begin{aligned} det[\mathbf{S}_1^{-1}(\mathbf{A}-\lambda\mathbf{I})\mathbf{S}_1] = det(\mathbf{A}-\lambda\mathbf{I}) &= det\begin{pmatrix} \lambda_1-\lambda & * \\ \mathbf{0} & \mathbf{B}-\lambda\mathbf{I} \end{pmatrix} \\ &= det(\lambda_1-\lambda)\cdot det(\mathbf{B}-\lambda\mathbf{I}), \end{aligned}$$

so that every eigenvalue of $\mathbf{A}$ is an eigenvalue of $\mathbf{B}$. The same argument can be applied to all steps of the process to prove that the diagonal elements of the resulting similarity transformation are all eigenvalues of the original matrix $\mathbf{A}$.

The process outlined above is a proof of the so-called Schur's theorem. This theorem can be stated as follows:

> Let $\mathbf{A}$ be an $n\times n$ matrix having only real eigenvalues. Then there is an $n\times n$ orthogonal matrix $\mathbf{S}$ such that $\mathbf{S}^T\mathbf{A}\mathbf{S}$ is an upper-triangular matrix.

Schur's theorem is more general than the statement above in that it holds when the matrix $\mathbf{A}$ has complex entries and complex eigenvalues. A more general statement of the theorem reads

> Given any matrix $\mathbf{A}$, there is a unitary matrix $\mathbf{S}$ such that $\mathbf{S}^H\mathbf{A}\mathbf{S}$ is an upper-triangular matrix.

Example 3. Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}.$$

Solution. To use Schur's theorem we first need to find one eigenvalue and its corresponding eigenvector. $\mathbf{A}$ has one eigenvalue $\lambda = 2$ and corresponding (normalized) eigenvector

$$\mathbf{x} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

We now construct an orthonormal basis $\{\mathbf{x},\ \mathbf{y}_1\}$, where

$$\mathbf{y}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Form the matrix $\mathbf{U} = [\mathbf{x}\ \ \mathbf{y}_1]$, or

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

with $\mathbf{U}^{-1} = \mathbf{U}^T$ and then

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{pmatrix} 2 & -2 \\ 0 & 2 \end{pmatrix}.$$

Example 4. Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} -2 & 1 & 0 \\ -4 & 0 & 3 \\ 0 & 1 & -2 \end{pmatrix}.$$

Solution. The matrix $\mathbf{A}$ has one eigenvalue $\lambda = -2$ and corresponding (normalized) eigenvector

$$\mathbf{x} = \frac{1}{5} \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}.$$

We now construct an orthonormal basis $\{\mathbf{x},\ \mathbf{y}_1,\ \mathbf{y}_2\}$, where

$$\mathbf{y}_1 = \frac{1}{5} \begin{pmatrix} 4 \\ 0 \\ -3 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Form the matrix $\mathbf{U} = [\mathbf{x}\ \ \mathbf{y}_1\ \ \mathbf{y}_2]$,

$$\mathbf{U} = \begin{pmatrix} 3/5 & 4/5 & 0 \\ 0 & 0 & 1 \\ 4/5 & -3/5 & 0 \end{pmatrix},$$

and then

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{pmatrix} -2 & 0 & 7/5 \\ 0 & -2 & 1/5 \\ 0 & -5 & 0 \end{pmatrix}.$$

Let $\mathbf{B}$ be the 2×2 submatrix in the lower-right-hand corner of $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$. Thus,

$$\mathbf{B} = \begin{pmatrix} -2 & 1/5 \\ -5 & 0 \end{pmatrix}.$$

An eigenvalue of $\mathbf{B}$ is $\lambda = -1$, with corresponding eigenvector

$$\mathbf{x} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}.$$

An orthonormal basis is $\{\mathbf{x}, \mathbf{y}_1\}$, where

$$\mathbf{x} = \frac{1}{\sqrt{26}}\begin{pmatrix} 1 \\ 5 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_1 = \frac{1}{\sqrt{26}}\begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

Construct the matrix

$$\mathbf{V} = \frac{1}{\sqrt{26}}\begin{pmatrix} 1 & 5 \\ 5 & -1 \end{pmatrix},$$

with

$$\begin{aligned} \mathbf{V}^{-1}\mathbf{B}\mathbf{V} &= \frac{1}{26}\begin{pmatrix} 1 & 5 \\ 5 & -1 \end{pmatrix}\begin{pmatrix} -2 & 1/5 \\ -5 & 0 \end{pmatrix}\begin{pmatrix} 1 & 5 \\ 5 & -1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & -26/5 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Now set

$$\mathbf{W} = \begin{pmatrix} 1 & | & 0 \\ - & | & - \\ 0 & | & \mathbf{V} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{26} & 5/\sqrt{26} \\ 0 & 5/\sqrt{26} & -1/\sqrt{26} \end{pmatrix}.$$

Then

$$\mathbf{U}\mathbf{W} = \frac{1}{5\sqrt{26}}\begin{pmatrix} 3\sqrt{26} & 4 & 20 \\ 0 & 25 & -5 \\ 4\sqrt{26} & -3 & -15 \end{pmatrix}$$

and

$$\mathbf{W}^{-1}(\mathbf{U}^{-1}\mathbf{A}\mathbf{U})\mathbf{W} = (\mathbf{U}\mathbf{W})^{-1}\,\mathbf{A}\,(\mathbf{U}\mathbf{W}) = \begin{pmatrix} -2 & 7/\sqrt{26} & -7/5\sqrt{26} \\ 0 & -1 & -26/5 \\ 0 & 0 & -1 \end{pmatrix}.$$

Example 5. Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 0 & 0 & -1 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 4 & 2 \\ -1 & 0 & 0 & 3 \end{pmatrix}.$$

Solution. To use Schur's theorem we first need to find one eigenvalue and its corresponding eigenvector. $\mathbf{A}$ has one eigenvalue $\lambda = 4$ and corresponding to the eigenvector

$$\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

We now construct a basis $\{\mathbf{x},\ \mathbf{y}_1,\ \mathbf{y}_2, \mathbf{y}_3\}$, where

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{y}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{y}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Form the matrix $\mathbf{U} = [\mathbf{x}\ \ \mathbf{y}_1\ \ \mathbf{y}_2]$,

$$\mathbf{U} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and then

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{pmatrix} 4 & 2 & 0 & 2 \\ 0 & 3 & 0 & -1 \\ 0 & 1 & 2 & 1 \\ 0 & -1 & 0 & 3 \end{pmatrix}.$$

Let $\mathbf{B}$ be the 3×3 submatrix in the lower-right-hand corner of $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$. Thus,

$$\mathbf{B} = \begin{pmatrix} 3 & 0 & -1 \\ 1 & 2 & 1 \\ -1 & 0 & 3 \end{pmatrix}.$$

An eigenvalue of $\mathbf{B}$ is $\lambda = 2$, with corresponding eigenvector

$$\mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

A basis is $\{\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2\}$, where

$$\mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \mathbf{y}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Construct the matrix

$$\mathbf{V} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

with

$$\mathbf{V}^{-1}\mathbf{B}\mathbf{V} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 0 & -1 \\ 1 & 2 & 1 \\ -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{pmatrix}.$$

Now set

$$\mathbf{W} = \begin{pmatrix} 1 & | & 0 \\ - & | & - \\ 0 & | & \mathbf{V} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\mathbf{W}^{-1}(\mathbf{U}^{-1}\mathbf{A}\mathbf{U})\mathbf{W} = (\mathbf{U}\mathbf{W})^{-1}\ \mathbf{A}\ (\mathbf{U}\mathbf{W}) = \begin{pmatrix} 4 & 0 & 2 & 2 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -1 & 3 \end{pmatrix}.$$

Let $\mathbf{C}$ be the 2×2 submatrix in the lower-right-hand corner of $\mathbf{U}\mathbf{W}^{-1}\mathbf{A}\mathbf{U}\mathbf{W}$. Thus,

$$\mathbf{B} = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

An eigenvalue of $\mathbf{C}$ is $\lambda = 2$, with corresponding eigenvector

$$\mathbf{x} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

A basis is $\{\mathbf{x}, \mathbf{y}_1\}$, where

$$\mathbf{x} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{y}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Construct the matrix

$$\mathbf{Y} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

with

$$\mathbf{Y}^{-1}\mathbf{C}\mathbf{Y} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Now set

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 & | & 0 \\ 0 & 1 & | & 0 \\ - & - & | & - \\ 0 & 0 & | & \mathbf{Y} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Then

$$\mathbf{Z}^{-1}(\mathbf{U}\mathbf{W}^{-1}\mathbf{A}\mathbf{U}\mathbf{W})\mathbf{Z} = (\mathbf{U}\mathbf{W}\mathbf{Z})^{-1}\ \mathbf{A}\ (\mathbf{U}\mathbf{W}\mathbf{Z}) = \begin{pmatrix} 4 & 0 & 2\sqrt{2} & 0 \\ 0 & 2 & \sqrt{2} & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Example 6. Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{pmatrix}.$$

Solution. To use Schur's theorem first need an to find one eigenvalue and its corresponding eigenvector. $\mathbf{A}$ has one eigenvalue $\lambda = 1$ and corresponding to the eigenvector

$$\mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

We now construct a basis $\{\mathbf{x},\ \mathbf{y}_1,\ \mathbf{y}_2\}$, where

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Form the matrix $\mathbf{U} = [\mathbf{x}\ \ \mathbf{y}_1\ \ \mathbf{y}_2]$,

$$\mathbf{U} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and then

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{pmatrix} 1 & -1 & -3/2 \\ 0 & 4 & 5/2 \\ 0 & -2 & -2 \end{pmatrix}.$$

Let $\mathbf{B}$ be the 2×2 submatrix in the lower-right-hand corner of $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$. Thus,

$$\mathbf{B} = \begin{pmatrix} 4 & 5/2 \\ -2 & -2 \end{pmatrix}.$$

An eigenvalue of $\mathbf{B}$ is $\lambda = -1$, with corresponding eigenvector

$$\mathbf{x} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

A basis is $\{\mathbf{x}, \mathbf{y}_1\}$, where

$$\mathbf{x} = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Construct the matrix

$$\mathbf{V} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix},$$

with

$$\mathbf{V}^{-1}\mathbf{B}\mathbf{V} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 4 & 5/2 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 9/2 \\ 0 & 3 \end{pmatrix}.$$

Now set

$$\mathbf{W} = \begin{pmatrix} 1 & | & 0 \\ - & | & - \\ 0 & | & \mathbf{V} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & -2 & 1 \end{pmatrix}.$$

Then

$$\mathbf{W}^{-1}(\mathbf{U}^{-1}\mathbf{A}\mathbf{U})\mathbf{W} = (\mathbf{U}\mathbf{W})^{-1}\ \mathbf{A}\ (\mathbf{U}\mathbf{W}) = \begin{pmatrix} 1 & 2 & -7/2 \\ 0 & -1 & 9/2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Example 7. Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & 2 \\ 0 & -1 & -1 \\ -1 & 1 & 2 \end{pmatrix}.$$

Solution. To use Schur's theorem first need an to find one eigenvalue and its corresponding eigenvector. $\mathbf{A}$ has one eigenvalue $\lambda = 0$ and corresponding to the eigenvector

$$\mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

We now construct a basis $\{\mathbf{x},\ \mathbf{y}_1,\ \mathbf{y}_2\}$, where

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_2 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}.$$

Form the matrix $\mathbf{U} = [\mathbf{x}\ \ \mathbf{y}_1\ \ \mathbf{y}_2]$,

$$\mathbf{U} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 1 & 0 & -2 \end{pmatrix},$$

and then

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{pmatrix} 0 & 1/3 & -5 \\ 0 & -1/2 & -3/2 \\ 0 & 1/6 & 1/2 \end{pmatrix}.$$

Let $\mathbf{B}$ be the 2×2 submatrix in the lower-right-hand corner of $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$. Thus,

$$\mathbf{B} = \frac{1}{6}\begin{pmatrix} -3 & -9 \\ 1 & 3 \end{pmatrix}.$$

An eigenvalue of $\mathbf{B}$ is $\lambda = 0$, with corresponding eigenvector

$$\mathbf{x} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

A basis is $\{\mathbf{x}, \mathbf{y}_1\}$, where

$$\mathbf{x} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Construct the matrix

$$\mathbf{V} = \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix},$$

with

$$\mathbf{V}^{-1}\mathbf{B}\mathbf{V} = \begin{pmatrix} -1/2 & -9/2 \\ 1/6 & 1/2 \end{pmatrix}^{-1} \begin{pmatrix} -1/2 & -3/2 \\ 1/6 & 1/2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & -1/6 \\ 0 & 0 \end{pmatrix}.$$

Now set

$$\mathbf{W} = \begin{pmatrix} 1 & | & 0 \\ - & | & - \\ 0 & | & \mathbf{V} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & -1 & 3 \end{pmatrix}.$$

Then

$$\mathbf{W}^{-1}(\mathbf{U}^{-1}\mathbf{A}\mathbf{U})\mathbf{W} = (\mathbf{U}\mathbf{W})^{-1}\ \mathbf{A}\ (\mathbf{U}\mathbf{W}) = \begin{pmatrix} 0 & 6 & -44/3 \\ 0 & 0 & -1/6 \\ 0 & 0 & 0 \end{pmatrix}.$$

Example 8. Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 5 & 2 & 1 \\ -8 & -3 & -2 \\ 7 & 4 & 3 \end{pmatrix}.$$

Solution. To use Schur's theorem first need an to find one eigenvalue and its corresponding eigenvector. $\mathbf{A}$ has one eigenvalue $\lambda = 1$ and corresponding to the eigenvector

$$\mathbf{x} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}.$$

We now construct a basis $\{\mathbf{x},\ \mathbf{y}_1,\ \mathbf{y}_2\}$, where

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}.$$

Form the matrix $\mathbf{U} = [\mathbf{x}\ \ \mathbf{y}_1\ \ \mathbf{y}_2]$,

$$\mathbf{U} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 2 & 0 & 1 \end{pmatrix},$$

and then

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{pmatrix} 1 & 22/5 & 6 \\ 0 & 5 & 5 \\ 0 & -9/5 & -1 \end{pmatrix}.$$

Let $\mathbf{B}$ be the 2×2 submatrix in the lower-right-hand corner of $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$. Thus,

$$\mathbf{B} = \begin{pmatrix} 5 & 5 \\ -9/5 & -1 \end{pmatrix}.$$

An eigenvalue of $\mathbf{B}$ is $\lambda = 2$, with corresponding eigenvector

$$\mathbf{x} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}.$$

A basis is $\{\mathbf{x}, \mathbf{y}_1\}$, where

$$\mathbf{x} = \begin{pmatrix} 5 \\ -3 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y}_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}.$$

Construct the matrix

$$\mathbf{V} = \begin{pmatrix} 5 & 3 \\ -3 & 5 \end{pmatrix},$$

with

$$\mathbf{V}^{-1}\mathbf{B}\mathbf{V} = \begin{pmatrix} 5 & 3 \\ -3 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 5 & 5 \\ -9/5 & -1 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ -3 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 28/5 \\ 0 & 2 \end{pmatrix}.$$

Now set

$$\mathbf{W} = \begin{pmatrix} 1 & | & 0 \\ - & | & - \\ 0 & | & \mathbf{V} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 3 \\ 0 & -3 & 5 \end{pmatrix}.$$

Then

$$\mathbf{W}^{-1}(\mathbf{U}^{-1}\mathbf{A}\mathbf{U})\mathbf{W} = (\mathbf{U}\mathbf{W})^{-1}\ \mathbf{A}\ (\mathbf{U}\mathbf{W}) = \begin{pmatrix} 1 & 4 & 43.2 \\ 0 & 2 & 6.8 \\ 0 & 0 & 2 \end{pmatrix}.$$

1) Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & -1 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

2) Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 0 & 1 \\ 1 & 3 & -1 \\ -1 & 0 & 2 \end{pmatrix}.$$

3) Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 4 & -1 & 1 \\ -1 & 4 & -1 \\ 1 & -1 & 4 \end{pmatrix}.$$

4) Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -2 \\ -1 & 1 & 0 \\ 0 & -2 & 4 \end{pmatrix}.$$

When a $n \times n$ matrix has n linearly independent eigenvectors, then Schur's theorem yields a diagonal matrix having the eigenvalues along the main diagonal. The triangularization work is simplified when the matrix $\mathbf{A}$ already has zero entries below the first element of the first column.

5) Use Schur's theorem to triangularize the matrix

$$\mathbf{A} = \begin{pmatrix} 5 & 3 & -2 \\ 0 & 2 & 6 \\ 0 & 1 & 3 \end{pmatrix}.$$

10 Gerschgorin's Theorem.

We have seen that to compute the eigenvalues of an $n \times n$ matrix we have to determine the characteristic equation of the matrix – which is a polynomial of degree n – and then find the roots of the characteristic equation. Since this task is obviously very difficult when n is large, some numerical methods have been developed to give information about the eigenvalues of a matrix.
Let $\mathbf{A}$ be an $n \times n$ matrix with real or complex entries

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

and denote the sum of the absolute values of the elements of the k^{th} row, except a_{kk}, to be

$$\rho_k = |a_{k,1}| + |a_{k,2}| + \dots + |a_{k,k-1}| + |a_{k,k+1}| + \dots + |a_{k,n}|$$

for $k = 1, 2, \dots, n$.
The circle in the complex plane centered at a_{kk} with radius ρ_k

$$\mathcal{D}_k = \{z \ : \ |z - a_{kk}| \leq \rho_k\}$$

is called a Gerschgorin's circle.

Example 6. The Gerschgorin's circles for the matrix

$$\mathbf{A} = \begin{pmatrix} 15 & 3+4i & 0 \\ -3 & 32i & 1-3i \\ 0.01 & -0.25 & 2-3i \end{pmatrix}$$

are

$$| z - 15 | \leq | 3 + 4i | = 5,$$

$$| z - 32i | \leq | -3 | + | 1 - 3i | = 6.162$$

$$| z - (2 - 3i) | \leq | 0.01 | + | -0.25 | = 0.26.$$

The following theorem due to Gerschgorin guarantees that the eigenvalues of an $n \times n$ matrix $\mathbf{A}$ are contained in the union of its n Gerschgorin's circles.

Gerschgorin's theorem. Let λ be an eigenvalue of an arbitrary $n \times n$ matrix $\mathbf{A} = (a_{jk})$. Then for some k, $1 \leq k \leq n$,

$$|a_{kk} - \lambda| \leq |a_{k,1}| + |a_{k,2}| + ... + |a_{k,k-1}| + |a_{k,k+1}| + ... + |a_{k,n}|. \quad (59)$$

For eack $k = 1, 2, ..., n$ the inequality (59) determines the corresponding Gerschgorin's circle of center a_{kk} and radius ρ_k in the complex plane. The theorem states that each of the eigenvalues of $\mathbf{A}$ lies in one of the n disks. The various disks in general overlap, but if one disk is disjoint from the remaining disks, then there must be an eigenvalue of $\mathbf{A}$ in that disk.

Proof. Let us prove that each eigenvalue of a square matrix $\mathbf{A}$ lies in at least one Gerschgorin's circle generated by $\mathbf{A}$. Let λ be the eigenvalue associated with the eigenvector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ . \\ x_k \\ . \\ x_n \end{pmatrix}. \quad (60)$$

Then,

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}. \quad (61)$$

Let x_k in (60) be the component of the eigenvector $\mathbf{x}$ that is the greatest in absolute value. Then we have

$$\left| \frac{x_j}{x_k} \right| \leq 1 \quad (62)$$

for $j = 1, 2, ..., n$.

The vector equation (61) is equivalent to a system of n equations for the n components of the vector $\mathbf{x}$. The k-component of that equation is

$$a_{k1}x_1 + a_{k2}x_2 + ... + a_{kk}x_k + ... + a_{kn}x_n = \lambda x_k$$

or, rearranging the terms

$$a_{kk} - \lambda = -a_{k1}\frac{x_1}{x_k} - ... - a_{k,k-1}\frac{x_{k-1}}{x_k} - a_{k,k+1}\frac{x_{k+1}}{x_k} - ... - a_{kn}\frac{x_n}{x_k}.$$

By taking absolute values on both sides of the above equation, using condition (63) and applying the triangle inequality

$$|a+b| \leq |a| + |b|,$$

we obtain

$$|\lambda - a_{kk}| \leq |a_{k1}| + ... + |a_{k,k-1}| + |a_{k,k+1}| + ... + |a_{kn}|,$$

and hence λ is in the k^{th} Gerschgorin's circle, which have center a_{kk} and radius

$$\rho_k = |a_{k1}| + ... + |a_{k,k-1}| + |a_{k,k+1}| + ... + |a_{kn}|.$$

Example 7. Use Gerschgorin's theorem to estimate the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 26 & -2 & 2 \\ 2 & 21 & 4 \\ 4 & 2 & 28 \end{pmatrix}.$$

Solution. From Gerschgorin's theorem it follows that the eigenvalues of this matrix lie in the three disks.

$$\begin{aligned} \mathcal{D}_1 &: |z-26| \leq |-2| + |2| = 4 \\ \mathcal{D}_2 &: |z-21| \leq |2| + |4| = 6 \\ \mathcal{D}_3 &: |z-28| \leq |2| + |4| = 6. \end{aligned}$$

Thus, one eigenvalue is in the circle centered at 26, radius 4; the second disk is centered at 21 and has radius 6; the third disk is centered at 28 and has radius 6. For the sake of comparison with the exact values, the eigenvalues of $\mathbf{A}$ are $30, 25$ and 20.

5) Use Gerschgorin's theorem to localize the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 5 & 1 & 1 \\ 0 & 6 & 1 \\ 1 & 0 & -5 \end{pmatrix}$$

and compare with the exact values.

6) Use Gerschgorin's theorem to localize the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 12 & 0 & -4 \\ 1 & 0 & -1 & 0 \\ 0 & 5 & 0 & 0 \end{pmatrix}$$

and compare with the exact values.

7) A matrix $\mathbf{A}$, where

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

is said to be *diagonally dominant* if

$$|a_{kk}| > \sum_{j \neq k}^{n} |a_{kj}|$$

for each $k = 1, 2, \dots n$.
Show that Gerschgorin's theorem guarantees that diagonally dominant matrices cannot have a zero eigenvalue.

11 Computer Exercises: Using MATLAB.

When you go to the computer lab, you will see a number of terminals. Sit down in front of any one of them. The screen will probably be blank. Touch the mouse and a blue screen will show up. On it you will see a login field. Enter your usename. The system will ask you your password. Enter it.
If your username and password are correct, another blue screen will show up. On the right side of this screen you will see a column displaying *netscape*, *mathematica*, *matlab*, *mapple* and three hosts named *local*, *line* and *plane*. Click on matlab and a matlab window will pop up.
If you start typing matlab commands right now you will not have a file to store your work. This is okay if you just want to do a quick computation and are not interested in saving your work.
If you do want to save your work, before you start typing matlab commands you need first type diary('johndoe1'). That **johndoe1** is any name you wish to give your file. To keep a neat record of your computer exercises, it is advisable to give the file you created your name, and number it in sequence according to the number of assignments we will have. Once you have a file called *johndoe1*, the input and output that follow will be saved in it. The command diary records everything that you do, including your mistakes.
When you have finished your work type **quit** to quit matlab. If you go to your home directory and type **more johndoe1**, you will see your work, input and output. If you want to edit your file to erase the mistakes, use any editor you know, such as **pico** or **emacs**, by typing **pico johndoe1** or **emacs johndoe1**. If you want to print what you have done type **lpr johndoe1**.
Let us try a few commands to see how things work. You do not need to save this.

Entering matrices in MATLAB is easy. To enter the matrix

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$$

type

$$\mathbf{A} = [1\ 2\ 3\ 4 \\ 5\ 6\ 7\ 8 \\ 9\ 10\ 11\ 12 \\ 13\ 14\ 15\ 16]$$

In this way we can create any arbitrary matrix. Try it yourself!
Matlab has also a number of special built-in functions for special matrices. Take a look at the various matrices that are generated with the following commands:

eye(4), **hadamard**(8), **toeplitz**([2,-1,zeros(1,4)]), **cyclic**(5), **consec**(4).

Operations on matrices are very easy as well. Type, for instance, $A = consec(4)$. Matlab will return the matrix you just typed before. To multiply matrices use $\mathbf{A} * \mathbf{B}$. For

The determinant of a matrix is $\mathbf{det}(\mathbf{A})$. The inverse is $\mathbf{inv}(\mathbf{A})$.

Before doing the problems of this assignment, practice the matlab with the examples in the book. You will learn more as you do those exercises. Try to familiarize yourself with Matlab as you go along solving the problems. When you feel at easy with matlab begin this assignment.

11.1 Matrices and determinants

1) We have seen that the three elementary row operations are

Type I. Interchange any two rows of **A**

Type II. Multiply any row of **A** by a nonzero constant

Type III. Add a constant multiple of one row to another row.

To interchange of rows i and j of a matrix **A** Matalb uses the command **rowswap(A,i,j)**. Set

n=5; i=2; j=4;
A=randint(n)
B=rowswap(**A**, i, j)
d1=det(**A**)
d2=det(**B**)

and compare the determinants **d1** and **d2**.

To multiply row i of a matrix **A** by a nonzero scalar c, Matalb uses the command **rowscale(A,i,c)**. Choose any c and set

n=4; i=3; j=5;
A=randint(n)
B=rowscale(**A**, i, c)
d1=det(**A**)
d2=det(**B**)

and compare the determinants **d1** and **d2**.

To add a multiple of one row of a matrix **A** to another row of **A**, Matalb uses the command **rowcomb(A,i,j,c)**, which adds **c** times row i to row j of **A**. Set

n=5; i=2; j=3; c=2;
A=randint(n)
B=rowcomb(**A**, i, j,c)
d1=det(**A**)
d2=det(**B**)

and compare the determinants **d1** and **d2**.

Make a conjecture about the effect of each of the three row operations on the value of the determinant of a matrix.

2) Let $\mathbf{B}_n$ be the tridiagonal matrix that have 5's on the main diagonal and 2's on the diagonals just above and below the main diagonal. For example, when $n = 4$ we have

$$\mathbf{B}_4 = \begin{pmatrix} 5 & 2 & 0 & 0 \\ 2 & 5 & 2 & 0 \\ 0 & 2 & 5 & 2 \\ 0 & 0 & 2 & 5 \end{pmatrix}.$$

a) For n=1, 2, 3, 4, 5, compute $det(\mathbf{B}_n)$.

b) Use Matlab's symbolic toolbox to find the determinant of the matrix

$$\mathbf{C} = \begin{pmatrix} c^2+1 & c & 0 \\ c & c^2+1 & c \\ 0 & c & c^2+1 \end{pmatrix},$$

where c is a real number.

c) Observe that the matrix **C** fits the pattern of the $\mathbf{B}_n$ matrices: there is a number c on the sub- and super-diagonals, $c^2 + 1$ on the main diagonal and zeros everywhere else. Use your result in **b)** to make a conjecture about the general expression for $det(\mathbf{B}_n)$ that you have calculated in **a)**.

d) Find a recursion relation for determining $det(\mathbf{B}_n)$.

e) Use the recursion relation (and induction on n) to prove that your conjecture is correct.

3) Form a random 3×4 of rank 2 matrix by using the command

A=randint(3,4,5,2).

Compute its reduced row echelon form by using the command **rref(A)**.

4) Consider the 5×5 matrix

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ x & 0 & 0 & 0 & 0 \end{pmatrix}$$

for all possible values of x.

a) For what values of x is $\mathbf{P}$ non-singular?

b) Use Matlab to compute $\mathbf{P}^{-1}$ for several values of x. Then guess the general formula for $\mathbf{P}^{-1}$ and prove that your formula is correct.

c) Use Matlab to compute $\mathbf{P}^5$ for several values of x. Then guess the general formula for $\mathbf{P}^5$ and prove that your formula is correct.

5) A square matrix is called *strictly upper triangular* if its entries on or below the main diagonal are all zero.

a) Use the Matlab command **triu(rand(4),1)** to generate 4×4 strictly upper triangular matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$. Then calculate the products $\mathbf{AB}$, $\mathbf{ABC}$, $\mathbf{ABCD}$. What pattern do you observe about the forms of these matrix products?

b) Use the Matlab command **triu(rand(4),1)** to generate a 4×4 strictly upper triangular matrix $\mathbf{A}$. Set $\mathbf{I} = eye(4)$, the 4×4 identity matrix. Calculate the matrix $\mathbf{U} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + ... + \mathbf{A}^k$ and the product $(\mathbf{I} - \mathbf{A})\mathbf{U}$ for $k = 3, 4, 5$. What do you observe?

c) Based on the familiar geometric series

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + ...,$$

which is valid for all real numbers x that satisfy $|x| < 1$, make a conjecture about the value of $(\mathbf{I} - \mathbf{A})\mathbf{U}$ for a suitable $\mathbf{A}$.

6) Create three matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ as follows:

A=randint(3,3,5,3)

B=randint(3,1,5)

C=randint(1,3,5)

a) Solve the system of three equations in three unknowns $\mathbf{Ax} = \mathbf{B}$ by using the operator $\mathbf{x} = \mathbf{A} \backslash \mathbf{B}$.

b) Use Matlab to solve the system $\mathbf{Ax} = \mathbf{B}$ by using $\mathbf{A}^{-1}$. Does your answer agree with the result obtained in **a)**?

c) Solve the system of three equations in three unknowns $\mathbf{yA} = \mathbf{C}$ by using the operator $\mathbf{y} = \mathbf{C}/\mathbf{A}$.

d) Solve the system $\mathbf{yA} = \mathbf{C}$ by converting it to the form $\mathbf{Ex} = \mathbf{F}$ through the use of transposes and then apply the same method as in **b)**. Does this answer agree with your answer in **c)**?

7) Let $\mathbf{A}_n$ be the matrix

$$\mathbf{A}_n = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 2 & 2 & 0 & 0 & \dots & 0 & 0 \\ 1 & 2 & 3 & 3 & 0 & \dots & 0 & 0 \\ 1 & 2 & 3 & 4 & 4 & \dots & 0 & 0 \\ . & . & . & . & . & \dots & . & . \\ 1 & 2 & 3 & 4 & 5 & \dots & (n-1) & (n-1) \\ 1 & 2 & 3 & 4 & 5 & \dots & (n-1) & n \end{pmatrix}.$$

a) Find the determinants of $\mathbf{A}_4$, $\mathbf{A}_6$, $\mathbf{A}_8$, $\mathbf{A}_{10}$ and $\mathbf{A}_{12}$. Make a conjecture about the value of the determinant of $\mathbf{A}_n$ for all values of n.

b) Create matrices $\mathbf{B}_4, \mathbf{B}_6$, and $\mathbf{B}_8$ that have the same elements as $\mathbf{A}_4$, $\mathbf{A}_6$ and $\mathbf{A}_8$ respectively, except for the element *zero* on the upper right corner that is replaced by 1. Would you expect the values of the determinants of the **B** matrices to be different from the value of the determinants of the corresponding **A** matrices? Ckeck out your answer by finding $det(\mathbf{B}_4)$, $det(\mathbf{B}_6)$, and $det(\mathbf{B}_8)$.

c) Would you expect the value of the determinants of the $\mathbf{B}$ matrices to be different from the value of the determinants of the $\mathbf{A}$ matrices if the zero entry on the upper right corner of the $\mathbf{A}$ matrices was replaced by a larger number, say 10, instead of 1? Ckeck out your answer by actual finding the determinants. Are the matrice $\mathbf{A}_n$ sensitive to changes in elements on the upper right corner?

d) Create matrices $\mathbf{C}_4, \mathbf{C}_6$, and $\mathbf{C}_8$ that have the same elements as $\mathbf{A}_4$, $\mathbf{A}_6$ and $\mathbf{A}_8$ respectively, except for the element *one* on the bottom left corner that is replaced by 0.9. Find $det(\mathbf{B}_4)$, $det(\mathbf{B}_6)$, and $det(\mathbf{B}_8)$. How does these changes affect each of the determinants?

d) Find the eigenvalues of $\mathbf{A}_{12}$.

12 Computer Exercises: Systems of equations, rank, span and subspaces.

1) Use MATLAB to solve the following system of equations

$$\begin{aligned} x + 2y + 3z &= 4 \\ 5x + 6y + 7z &= 8 \\ 9x + 10z + 11z &= 12. \end{aligned}$$

a) Write the system as a matrix equation $\mathbf{Ax} = \mathbf{B}$ and use MATLAB command $\mathbf{x} = \mathbf{A} \backslash \mathbf{B}$ to find a solution to the system. Does MATLAB give you the general solution to this system?

b) Use the command **rank(A)** to find the rank of the coefficient matrix of the system and the rank of the augmented matrix of the system. What does this tell you about the existence of solutions?

c) Use the command **rref(A)** to find the reduce row-echelon form of the augmented matrix of the system. From this result, find the general solution of the system. Does your general solution contain MATLAB's?

2) Given the system of two equations

$$\begin{aligned} 4x + y &= 7 \\ 3x - 2y &= -3, \end{aligned}$$

the MATLAB command **plotline** graphs each of the equations in this system.

a) Type **plotline(4,1,7)**, where the numbers 4,1 and 7 are the coefficients of the first equation in the system above.

b) If you want to add the second line of the system of equations to the existing picture, type first the command **hold on** so that MATLAB will freeze the current plot. Now enter the second line with **plotline(3, -2 ,-3)**.

c) From the geometry of the system you see on the screen, determine how many solutions the system has.

d) Now that you know how many solutions the above system has, let us estimate the values of x and y. Type **grid** to display a grid on your plot. Read off the values for x and y. Check your answer.

e) For each of the following systems, determine whether there is no solution, a unique solution – and, in this case estimate it graphically with the help of **grid** –, or many solutions. Before entering a new system, type the command **hold off** to unfreeze the previous plot. Then when you type plotline again for the new system the previous plot will disappear. If you want to save the previous plot in a file, click on the field *file* on the plot window and save it. Remember to type hold on after each plotline command.

i)

$$\begin{aligned} 4x + y &= 7 \\ -8x - 2y &= -4. \end{aligned}$$

Systems with more equations than unknowns are often assumed to be inconsistent. What say you about the systems below?

ii)

$$\begin{aligned} 4x + y &= 7 \\ 3x - 2y &= -3 \\ x + 3y &= -1. \end{aligned}$$

iii)

$$\begin{aligned} x + 2y &= 2 \\ x - 8y &= -3 \\ 3x - 4y &= 1 \\ 17x + 2y &= 18. \end{aligned}$$

iv)

$$\begin{aligned} x + y &= 3 \\ 3x - 2y &= 4 \\ x + 3y &= 9 \end{aligned}$$

3) This exercise refers to the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 5 & -2 & 1 \\ -1 & 2 & 3 & 2 & 3 \\ 4 & -1 & 2 & -1 & 2 \\ 2 & -1 & 0 & 3 & 4 \end{pmatrix}.$$

a) Use the command **rank(A)** to find rank of the matrix.

b) Use the command **rref(A)** to find the reduce row-echelon form of **A**. Let us denote the reduce row-echelon form by **B**.

c) Explain how to find a basis for the row space of **A** directly from the matrix **B**. State the dimension of the row space.

d) What relationship do you think exists between the rank, the number of nonzeros row of **B** and the number of vectors in the row space of **A**?

e) By hand, express each row of **A** as a linear combination of the basis vector for the row space.

f) By hand, solve the system $\mathbf{Bx} = \mathbf{0}$ and find a basis for the null space of **A**. State the dimension of the null space.

g) Use the command **nulbasis(A)** to find the null space of **A**. Does your result in **f)** agree with MATLAB's?

4) Let $\mathbf{A}$ be the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & -4 \\ -3 & 6 \end{pmatrix}.$$

a) By hand, find a basis $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ for the row space of $\mathbf{A}$.

b) By hand, find a basis $\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ for the null space of $\mathbf{A}$.

c) In MATLAB, type the command **drawvec**$([x_1, x_2])$, where x_1, x_2 are the values of the coordinates of the vector $\mathbf{X}$ you obtained above. What happens? Type **hold on** before proceeding.

d) Type **drawvec**$([y_1, y_2], 'g')$, where y_1, y_2 are the values of the coordinates of the vector $\mathbf{Y}$ you obtained above. Is it cool or not? What (geometric) relationship do you observe between the vectors $\mathbf{X}$ and $\mathbf{Y}$?

e) Type the command **figure(2)**. WOW!! Repeat the steps **a), b), c), d)** for the matrix

$$\begin{pmatrix} -1 & 3 \\ 4 & -12 \end{pmatrix}.$$

f) From what you observe in these two examples, make a conjecture about the relationship between the vectors in the row space and the null space of 2×2 matrices of rank 1.

g) Try to confirm your conjecture by repeating the steps **a), b), c)** and **d)** for your own 2×2 matrix of rank 1. Before starting, type **figure(3)** so that a new graphics windows will pop up.

5) Let $\mathcal{W} = span\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 4 \\ 3 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 3 \\ 5 \end{pmatrix}.$$

a) In this exercise you are going to determine whether the vector

$$\mathbf{v} = \begin{pmatrix} 5 \\ 3 \\ -6 \\ 4 \\ -6 \end{pmatrix}$$

is in $\mathcal{W}$.

Before you start, however, you should check that the vectors in $\mathcal{W}$ are indeed linearly independent. Create a matrix $\mathbf{A} = [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \mathbf{v}_3]$ having as columns the vectors in $\mathcal{W}$. Type the command $\mathbf{rref}([\mathbf{A}, \mathbf{zeros}(\mathbf{5}, \mathbf{1})])$. This command converts the augmented matrix of the homogeneous system $\mathbf{Ax} = \mathbf{0}$ to reduced row-echelon form. From the reduced row-echelon form you have obtained explain why the vectors are linearly independent.

b) Type **format rat** so that MATLAB will treat rational numbers as fractions instead of decimals.

c) Enter the vector $\mathbf{v}$. Type the command $\mathbf{rref}([\mathbf{A}, \mathbf{v}])$.

d) From what you see on the screen determine whether the vector $\mathbf{v}$ is in $\mathcal{W}$. Explain your reasoning. If you conclude that $\mathbf{v}$ is in $\mathcal{W}$, give an explicit expression for $\mathbf{v}$ as a linear combination $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ of the vectors in $\mathcal{W}$.

e) Repeat this exercises by choosing for $\mathbf{v}_2$ the vector

$$\mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 4 \\ 3 \\ 2 \end{pmatrix}.$$

Is $\mathbf{v}$ in $\mathcal{W}$ now?

f) Repeat this exercise with an arbitrary vector $\mathbf{v}$ that you will choose yourself. Pick the first one that comes to your mind. Is it in $\mathcal{W}$?

13 Computer Exercises: Orthogonality.

1) In this exercise you are going to play the *coordinate game*. In it, you are given two vectors $\mathbf{u}$ and $\mathbf{v}$ in R^2 and you have to guess the values of scalars α and β so that the tip of the vector $\alpha\mathbf{u} + \beta\mathbf{v}$ will lie at some predetermined point.

To start the game, type the MATLAB command **cogame**. A graphic window will pop up. On it, click on **level 1** and you will see the vectors $\mathbf{u}$ and $\mathbf{v}$. A target point O is depicted as a blue circle on the screen. On the field **enter components** you will type the values of α and β that you think will make the tip of the vector $\alpha\mathbf{u} + \beta\mathbf{v}$ lie in the circle O.

a) Play this game many times until you find the best choice for α and β. Once you have the best choice, type **grid** to display a grid on your figure, and use the grid to determine the approximate value of the coordinates of the blue circle. Compare it with the value of $\alpha\mathbf{u} + \beta\mathbf{v}$ for your best choice of α and β.

b) Start this game again. With the vectors $\mathbf{u}$ and $\mathbf{v}$ that you are given on clicking **level 1**, find, by hand, an exact expression for α and β such that the tip of the vector $\alpha\mathbf{u} + \beta\mathbf{v}$ is at the point $P = (1.5, -2)$.

c) Without changing the vector $\mathbf{u}$ and the position of the target circle, how would you replace the vector $\mathbf{v}$ that you are given in **b)** so as to make it impossible to reach the target circle?

2) Suppose $\mathcal{V}$ is a k-dimensional subspace of R^n and let $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ be an ordered basis for $\mathcal{V}$. By ordered basis we mean that the vectors in $\mathcal{B}_V$ have been assigned a first position, a second position, and so on. In the present case $\mathbf{v}_1$ is the first vector, $\mathbf{v}_2$ is the second vector and $\mathbf{v}_k$ is the last vector. Any vector $\mathbf{x}$ in $\mathcal{V}$ can be wrriten as a linear combination of the basis vectors

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_k\mathbf{v}_k, \tag{63}$$

where the coefficients $x_1, x_2, ..., x_k$ are called the coordinates of $\mathbf{x}$ with respect to the basis $\mathcal{B}_V$. Equation (63) can be written in matrix form as

$$\mathbf{x} = \begin{pmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_k \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_k \end{pmatrix} = \mathbf{V}\mathbf{x}_{\mathcal{V}}, \tag{64}$$

where the vector

$$\mathbf{x}_{\mathcal{V}} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_k \end{pmatrix}$$

is called the *coordinate vector* of $\mathbf{x}$ with respect to the basis $\mathcal{B}_V$.

Let $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a basis of R^3, where

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 2 \\ -3 \end{pmatrix}.$$

a) Describe how to use MATLAB to verify that the above vectors are linearly independent. You should explain your procedure and give the MATLAB command you have used.

b) Use MATLAB to find the coordinates of the vector

$$\mathbf{x} = \begin{pmatrix} -3 \\ 1 \\ 5 \end{pmatrix}$$

with respect to the basis $\mathcal{B}_V$. Here again you have to figure out how to go about doing this using MATLAB. Explain your procedure and give the MATLAB command you have used.

c) From your computation in **b)** create the 1×3 coordinate vector $\mathbf{x}_{\mathcal{V}}$. Next, create a 3×3 matrix $\mathbf{V}$ whose columns are $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. Verify equation (64).

Since there are infinitely many bases for R^3, there are inifinitely many coordinates vectors for each given vector $\mathbf{x}$. To see this, let $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ be a second basis of R^3, where the vectors

$$\mathbf{w}_1 = \begin{pmatrix} -3 \\ 1 \\ 5 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 4 \\ -2 \\ -7 \end{pmatrix}, \qquad \mathbf{w}_3 = \begin{pmatrix} 8 \\ 1 \\ -11 \end{pmatrix}$$

are of course linearly independent.

d) Use MATLAB to find the coordinates of the vector $\mathbf{x}$ relative to the basis $\mathcal{B}_W$.

e) From your computation in **d)** create the 1×3 coordinate vector $\mathbf{x}_{\mathcal{W}}$, and a 3×3 matrix $\mathbf{W}$ whose columns are $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. Verify equation (64).

f) Use MATLAB to find the following expressions

$$\mathbf{V}^{-1}\mathbf{W}\mathbf{x}_{\mathcal{W}}, \qquad \mathbf{W}^{-1}\mathbf{V}\mathbf{x}_{\mathcal{V}}.$$

What do you observe? Is there anything strange about the result of the computation $\mathbf{W}^{-1}\mathbf{V}\mathbf{x}_v$? Give an analytic explanation why these expressions yield the results you observe.

3) a) Use the MATLAB command **rref** to find the reduced row-echelon form of the following matrices

$$\mathbf{A}_{4\times 5} = \begin{pmatrix} 1 & -3 & -5 & 4 & 1 \\ 0 & 1 & 1 & -1 & 3 \\ -1 & 3 & 5 & -4 & 2 \\ -2 & -4 & 0 & 2 & -1 \end{pmatrix}, \qquad \mathbf{A}_{6\times 3} = \begin{pmatrix} 1 & 5 & 1 \\ -2 & 0 & 3 \\ -1 & -3 & 0 \\ 3 & 5 & -2 \\ 1 & 2 & -1 \\ 0 & 3 & 2 \end{pmatrix}$$

$$\mathbf{A}_{4\times 4} = \begin{pmatrix} 3 & 4 & -6 & 4 \\ -4 & 0 & 0 & 8 \\ -3 & -2 & 3 & 1 \\ 2 & 6 & -9 & 11 \end{pmatrix}, \qquad \mathbf{A}_{3\times 7} = \begin{pmatrix} 1 & 2 & 5 & 8 & -1 & -2 & -3 \\ 0 & 1 & 1 & -1 & 3 & 7 & 3 \\ 2 & 5 & 11 & 15 & 1 & 3 & -3 \end{pmatrix}.$$

From this computation, find the dimension of the row and column spaces for each of the matrices. What do you observe? Make a conjecture about the relatonship between the dimensions of the column and row spaces of matrices.

b) Note that if $\mathbf{A}$ is $m \times n$ the row vectors of $\mathbf{A}$ are in R^n while the column vectors are in R^m. Does this observation change in any way your conjecture in **a**)?

c) Test your conjecture by creating your own matrix $\mathbf{A}$. What do you observe? How many leading ones do you find in each reduced row-echelon form?

d) Try to sketch a proof of your conjecture by paying attention to the number of leading ones in the reduced row-echelon forms of the matrices above.

4) If $\mathbf{Q}$ is a square matrix and the column vectors of $\mathbf{Q}$ form an orthonormal set, then we say that $\mathbf{Q}$ is an orthogonal matrix.

a) In MATLAB, generate the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ by setting

$$\mathbf{A} = eye(4) - 0.5 * ones(4), \qquad\qquad \mathbf{B} = cyclic(4)$$

and

$$\mathbf{C} = \frac{1}{100}\begin{pmatrix} 54 & -82 & 18 & 6 \\ 54 & 18 & -82 & 6 \\ 18 & 6 & 6 & -98 \\ 62 & 54 & 54 & 18 \end{pmatrix}.$$

b) Verify that $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are orthogonal matrices.

c) Use the command **inv** to compute the inverse of each of the three matrices. How are the inverses related to the original matrices?

d) Compute the products $\mathbf{AB}$, $\mathbf{BC}$ and $\mathbf{AC}$ and test to see if each product is an orthogonal matrix.

e) Set the column vector $\mathbf{x} = \mathbf{randint(4, 1)}$ and compute its length using the command $\mathbf{norm}(\mathbf{x})$. Next compute the lengths of $\mathbf{Ax}$, $\mathbf{Bx}$ and $\mathbf{Cx}$. What do you observe?

f) Prove that if $\mathbf{Q}$ is an $n \times n$ orthogonal matrix and $\mathbf{x} \in R^n$, then $||\mathbf{Qx}||^2 = ||\mathbf{x}||^2$, and hence $\mathbf{Qx}$ has the same length as $\mathbf{x}$. Remember that

$$\langle \mathbf{Qx}, \mathbf{Qx} \rangle = (\mathbf{Qx})^T \mathbf{Qx} = ||\mathbf{Qx}||^2.$$

5) The span of the column vectors of a matrix $\mathbf{A}$ is the column space of $\mathbf{A}$. In many applications where a matrix $\mathbf{A}$ does not have orthonormal column vectors, it is desirable to find an orthonormal basis for the column space. In this exercise you will find orthogonal bases for the column space of a matrix using four different methods. Since you will be comparing the orthogonal basis obtained with these various methods, it is advisable to perform your calculation using first the **format rat** and then the **format long** so that you will be able to see which method gives a better basis.
In MATLAB, enter the vectors

$$\mathbf{q}_1 = \begin{pmatrix} -149 \\ 537 \\ -27 \\ 122 \end{pmatrix}, \qquad \mathbf{q}_2 = \begin{pmatrix} -50 \\ 180 \\ -9 \\ 41 \end{pmatrix}, \qquad \mathbf{q}_3 = \begin{pmatrix} -154 \\ 546 \\ -25 \\ 129 \end{pmatrix}$$

as columns of a matrix $\mathbf{A}$.

a) The MATLAB command $\mathbf{orth}(\mathbf{A})$ produces an orthonormal basis for the column space of $\mathbf{A}$. Set $\mathbf{U} = \mathbf{orth}(\mathbf{A})$ and compute the product $\mathbf{U}^T\mathbf{U}$. From the result of your computation $\mathbf{U}^T\mathbf{U}$, explain why the columns of $\mathbf{U}$ constitute an orthonormal set of vectors.

b) The MATLAB command $[\mathbf{Q}, \mathbf{R}] = qr(\mathbf{A})$ produces an orthogonal matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{A}$ such that $\mathbf{A}_{m\times n} = \mathbf{Q}_{m\times m}\mathbf{R}_{m\times n}$. If $\mathbf{A}$ has rank r, then the first r columns of $\mathbf{Q}$ form an orthogonal basis for the column space of $\mathbf{A}$. To extract the desired columns of $\mathbf{Q}$ that form a basis for the column space of $\mathbf{A}$, use the command

$$\mathbf{Q} = \mathbf{Q}(:, 1 : r).$$

Compute $\mathbf{Q}^T\mathbf{Q}$. Do the columns of $\mathbf{Q}$ form an orthonormal set of vectors?

c) Let us now compute an orthogonal basis for the column space of $\mathbf{A}$ by using the classical version of the Gram-Schmidt process. Set

$$\mathbf{G} = gschmidt(\mathbf{A}, 1).$$

The columns of $\mathbf{G}$ form a basis for the column space of $\mathbf{A}$. Compute $\mathbf{G}^T\mathbf{G}$. Do the columns of $\mathbf{G}$ form an orthonormal set of vectors?

d) The classical version of the Gram-Schmidt process, as specified in most linear algebra textbooks, is prone to numerical instabilities. Because of this, MATLAB uses a modified Gram-Schmidt process that has been devised to display better numerical stability. Let us compute an orthogonal basis for the column space of $\mathbf{A}$ by using this modified Gram-Schmidt process. Set

$$\mathbf{S} = gschmidt(\mathbf{A}).$$

and compute $\mathbf{S}^T\mathbf{S}$. Do the columns of $\mathbf{S}$ form an orthonormal set of vectors?

e) Compare the products $\mathbf{U}^T\mathbf{U}$, $\mathbf{Q}^T\mathbf{Q}$, $\mathbf{G}^T\mathbf{G}$ and $\mathbf{S}^T\mathbf{S}$. Which of the four methods ($\mathbf{orth}, [\mathbf{Q}, \mathbf{R}], \mathbf{gschmidt}(\mathbf{A}, \mathbf{1})$ or $\mathbf{gschmidt}(\mathbf{A})$) does a better job of producing an orthonormal basis? Which method does the worst job?

14 Computer Exercises: The least-squares problem, eigenvalues and the Gerschgorin's Theorem.

1) A company's record shows that the amount spent on advertising (x) and the total sales (y) for each of the first six months 1999 are as reported in the following table.

advertising costs (x)	2.4	3.6	3.6	4.1	4.7	5.3
sales (y)	33.8	34.7	35.5	36.0	37.5	38.1

a) Enter the x and y coordinates in MATLAB as columns vectors $\mathbf{x}$ and $\mathbf{y}$. To compute the coefficient of the best linear least-squares fit to the data, use the commands

$$\mathbf{A} = [\mathbf{x}, ones(size(\mathbf{x}))] \qquad \text{and} \qquad \mathbf{c} = \mathbf{A} \backslash \mathbf{y}$$

b) Use the following command to plot both the linear function and the data points on the same axis system

$$\mathbf{t = 2 : 0.1 : 6};$$

$$\mathbf{z = polyval(c, t)};$$

$$\mathbf{plot(x, y, \ 'x' \ , t, z)}$$

c) Suppose that for the last six months of 1999 the actual sales of the company are as given in the table below.

advertising costs (x)	6.0	6.5	7.0	7.5	8.0	8.5
sales (y)	38.6	38.7	38.5	37.6	37.2	36.4

Augment the vectors $\mathbf{x}$ and $\mathbf{y}$ of item **a)** above to include these additional coordinates for the second half of the year. Plot the 12 points using the command $\mathbf{plot(x, y, 'x')}$. Remember to type **hold on** before continuing. From this graph, how well does the linear model in item **b)** predict sales for the second half of the year?

d) When the linear least-squares fit is not accurate enough, we might consider fitting a quadratic function $y = c_1x^2 + c_2x + c_3$ to the entire set of 12 data points. In this case, we seek a least-squares solution of the matrix equation

$$\begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ . & . & . \\ x_{12}^2 & x_{12} & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_{12} \end{pmatrix}.$$

To use MATLAB to compute the least-squares fit, set

$$\text{and} \qquad \mathbf{d = A \backslash y},$$

where $\mathbf{x}$ and $\mathbf{y}$ are the augmented vectors of item **c)**.

f) Type **hold on** before continuing. Use the following command to plot both the linear function and the data points on the same axis system

$$t = 2 : 0.1 : 12;$$

$$\mathbf{z} = \mathbf{polyvac(d, t)}$$

$$plot(\mathbf{x}, \mathbf{y}, \ '\mathrm{x}' \ , t, \mathbf{z})$$

g) Is the quadratic function a better fit to the 12 points of the data than the linear function fit?

2) If $\mathbf{A}$ is a 2×2 matrix and $\mathbf{x}$ is any unit vector, then the directions of $\mathbf{Ax}$ and $\mathbf{x}$ will generally differ. However, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ belonging to a real eigenvalue λ, then $\mathbf{Ax}$ will be in the same direction as either $\mathbf{x}$ or - $\mathbf{x}$, depending on whether λ is positive or negative.

a) Set $\mathbf{A} = [2 \quad 0; 1.5 \quad 0.5]$. Use the command **eigshow(A)** to run an animation showing how the vectors $\mathbf{x}$ and $\mathbf{Ax}$ change as $\mathbf{x}$ moves around the unit circle. To move the vector $\mathbf{x}$ around the circle, hold the left button of the mouse down and move the mouse slowly in a counterclockwise circle. From what you observe, answer the following questions.

b) How many times did $\mathbf{x}$ and $\mathbf{Ax}$ line up? What is the significance of this in terms of eigenvalues and eigenvectors of $\mathbf{A}$?

c) What type of eigenvalues does $\mathbf{A}$ have? Are they real or complex? Are they distinct or multiple? If the eigenvalues are real, how many are positive, negative or zero? Is $\mathbf{A}$ singular? Explain.

d) Type the command **grid** to display a grid on the screen. Give a rough estimate of the eigenvalues and eigenvectors of $\mathbf{A}$.

e) Compute the exact eigenvalues and eigenvectors of **A**. Do your answers in items **b)** and **c)** agree with your exact computation?

f) Repeat this exercise for each of the follwoing matrices

$$\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}, \qquad \begin{pmatrix} -1 & 2 \\ -2 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

3) For each n from 3 to 5, generate a random $n \times n$ matrix **A** with integer entries by setting $\mathbf{A} = \mathbf{randint(n)}$.

a) The characteristic polynomial of **A** should have integer coefficients. Compute the coefficients by setting $\mathbf{p} = \mathbf{round(poly(A))}$.

b) Use the MATLAB commands **trace**(**A**) and **det**(**A**) to compute the trace and the determinant of **A**. What relations, if any, do you observe betweeen the coefficients and the values of the trace and determinant?

c) Compute the eigenvalue of **A** by setting $\mathbf{e} = \mathbf{eig(A)}$. Use MATLAB functions **sum**(**e**) and **prod**(**e**) to compute the sum and the product of the eigenvalues. How are these numbers related to the coefficients of the characteristic polynomials? How do they relate to the trace and determinant of **A**.

d) The MATLAB command **polyvalm** can be used to evaluate the polynomial of a matrix. For example, if $p(x) = x^2 + x + 2$ and **B** is a square matrix, then $p(\mathbf{B}) = \mathbf{B}^2 + \mathbf{B} + 2\mathbf{I}$. If the coefficients of $p(x)$ are stored in the vector $\mathbf{p} = [1, 1, 2]$, then the command **polyvalm**(**p**, **B**) will generate the matrix $p(\mathbf{B})$. For each of the random integer matrix **A** evaluate its characteristic polynomial $p(\mathbf{A})$ using the command **polyvalm**(**p**, **A**). What do you observe? Explain your result.

4) MATLAB's *gallery* matrices are matrices whose eigenvalues are nice even though they are difficult to compute accurately. Construct the 5×5 gallery matrix by setting $\mathbf{A} = \mathbf{gallery(5)}$.

a) Compute the eigenvalues of $\mathbf{A}$ be setting $\mathbf{e} = \mathbf{eig}(\mathbf{A})$. Compute also $\mathbf{A}^5$. What are the eigenvalues of $\mathbf{A}^5$? In general how would you expect the eigenvalues of $\mathbf{A}^5$ to be related to those of $\mathbf{A}$?

b) Let us now compute the eigenvalues in a different way by finding the roots of the characteristic polynomial. Compute the coefficients of the characteristic polynomial by setting $\mathbf{p} = \mathbf{poly}(\mathbf{A})$. Use **format long** and the command **roots(p)** to compute the roots of the characteristic polynomial. How do these compare to the eigenvalues computed in **a)**?

c) Explain why the characteristic polynomial of $\mathbf{A}$ must have integer coefficients. Set $\mathbf{p} = \mathbf{round}(\mathbf{p})$ and compute the roots again. Compare your results to those in **a)** and **b)**.

d) To understand what is going on, let us find the characteristic polynomial in functional form. The command $\mathbf{p} = \mathbf{poly}(\mathbf{sym}(\mathbf{A}))$ will return a polynomial of fifth degree in the variable x. Find the roots of this polynomial by setting $\mathbf{solve}(\mathbf{p})$. What do you conclude now?

e) A matrix $\mathbf{B}$ is said to be *nilpotent* if $\mathbf{B}^k = 0$ for some positive integer k. Prove that the eigenvalues of a nilpotent matrix must all be zero.

5) In 1933 S. Gerschgorin proved two theorems that give useful information about the location in the complex plane of the eigenvalues of a matrix without actually having to compute the eigenvalues. The first Gerschgorin's theorem states that

> *Every eigenvalue of an $n \times n$ matrix $\mathbf{A}$ lies in at least one of the circles having centers $\rho_j = a_{jj}$ and radii $r_j = \sum_{k \neq j} |a_{jk}|$, $j = 1, 2, ..., n$.*

MATLAB plots the Gerschgorin's circles of a matrix $\mathbf{A}$ using the command $\mathbf{gersch}(\mathbf{A})$. The command $\mathbf{gersch}(\mathbf{A}, \mathbf{1})$ will plot both the Gerschgorin's circles and the eigenvalues of $\mathbf{A}$.

a) Given the following matrices

$$\begin{pmatrix} 10 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & 1 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 10 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 10 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & -2 \end{pmatrix},$$

$$\begin{pmatrix} 10 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & 6 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 10 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & 5 & -15 \end{pmatrix}, \qquad \begin{pmatrix} 10 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & 15 & -20 \end{pmatrix},$$

plot the Gerschgorin's circles and the eigenvalues of each of the matrices above. Must each circle contain at least one eigenvalue? Is it possible that an eigenvalue lies *on* the circle as opposed to *inside* the circle?

b) The second Gerschgorin's theorem states that

> *If k of the Gerschgoring's circles of* $\mathbf{A}$ *are disjoint from the other circles, then there are k eigenvalues of* $\mathbf{A}$ *lying inside the union of those circles.*

Does this happen to any eigenvalue of the matrices in item **a)**?

15 A few extra problems.

1) Find the LU-decomposition for the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 1 \\ 1 & 4 \end{pmatrix}.$$

2) Let

$$A = \begin{pmatrix} a & b & c \\ 1 & 1 & 1 \end{pmatrix}.$$

Find conditions on a,b and c such that

a) rank(A)= 1 **b)** rank(A)=2.

3) Find the determinant

$$\begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{vmatrix}.$$

4) Solve the linear system

$$\begin{aligned} x + 2y - 3z &= 4 \\ x + 3y + z &= 11 \\ 2x + 5y - 4z &= 13 \\ 2x + 6y + 2z &= 22 \end{aligned}$$

5) Find the inverse of

$$\begin{pmatrix} \mathbf{I} & \mathbf{X} & \mathbf{Y} \\ \mathbf{0} & \mathbf{I} & \mathbf{Z} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

6) Find a basis for the row, column and null spaces of the following matrices

$$\begin{pmatrix} 1 & 1 & 5 & 1 & 4 \\ 2 & -1 & 1 & 2 & 2 \\ 3 & 0 & 6 & 0 & -3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 & 2 & 0 & 1 \\ 2 & 2 & 5 & 0 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 8 & 11 & 19 & 0 & 11 \end{pmatrix}.$$

7) Find a basis for the subspace S of R^3 defined by the equation $x+2y+3z=0$.

8) Find a basis for the null space of

$$\mathbf{A} = \begin{pmatrix} 3 & -1 & -1 & -3 \\ -2 & 2 & -2 & 2 \\ -1 & -1 & 3 & 1 \end{pmatrix}.$$

9) Find a basis for the null space of

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 4 & 3 & -2 \\ 7 & 4 & -2 & 6 & 1 \end{pmatrix}.$$

10) Peaches are 30 cents each, pears are 20 cents each, apples are 35 cents each and grapefruits are 50 cents each. Anne wants to get 5 peaches, 3 pears, 2 apples and 2 grapefruits. Form the price vector $\mathbf{p}$ whose *coordinates* are the prices of the respective fruits, and the demand vector $\mathbf{d}$ whose coordinates are the quantities of each fruit. Find and interpret the scalar product $\mathbf{p} \cdot \mathbf{d}$.

An elementary row operation on a matrix $\mathbf{A}$ is any one of the following.

Type I. Interchange any two rows of $\mathbf{A}$

Type II. Multiply any row of $\mathbf{A}$ by a nonzero constant

Type III. Add a constant multiple of one row to another row.

11) Put the matrix
$$\mathbf{A} = \begin{pmatrix} 2 & 3 & -1 \\ 4 & -1 & 5 \\ 0 & 2 & 6 \end{pmatrix}$$
in the form
$$\mathbf{A} = \begin{pmatrix} 1 & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$
using no more than three elementary row transformations.

12) Let $A = \begin{pmatrix} 1 & 4 \\ -3 & 1 \end{pmatrix}$. Show that A is non-singular and find A^{-1}.

13) Find a matrix P such that $PA = B$, where
$$A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 2 \\ 3 & 1 & 5 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & -1 & 2 \\ 1 & -2 & 3 \\ 3 & 1 & 5 \end{pmatrix}.$$

14) Show that $A = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}$ is singular and find a non-singular matrix P such that PA has last row zero.

15) Verify that the matrix $\frac{1}{6}\begin{pmatrix} 3 & 3 & 3 & 3 \\ 3 & -5 & 1 & 1 \\ 3 & 1 & -5 & 1 \\ 3 & 1 & 1 & -5 \end{pmatrix}$ is orthogonal.

16) Colleges x, y and z offer various programs, as shown in matrix **A** below:

$$\begin{array}{l} \\ \text{College x} \\ \text{College y} \\ \text{College z} \end{array} \begin{array}{c} \begin{array}{ccc} \text{forestry} & \text{engineering} & \text{music} \end{array} \\ \left| \begin{array}{ccc} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{array} \right| \end{array}$$

The programs, in turn, require various courses, as shown in matrix **B** below:

$$\begin{array}{c|ccc|} & \text{calculus} & \text{botany} & \text{harmony} & \\ \text{forestry} & 1 & 1 & 1 \\ \text{engineering} & 1 & 0 & 1 \\ \text{music} & 0 & 1 & 1 \end{array}.$$

Form the product **AB** and deduce which of the courses must be offered by each college.

17) Given an $n \times n$ matrix $\mathbf{A}$, we define the exponential $e^{\mathbf{A}}$ as

$$e^{\mathbf{A}} = \mathbf{I}_n + \sum_{k \geq 1} \frac{\mathbf{A}^k}{k!}$$

where $\mathbf{I}_n$ denotes the $n \times n$ indentity matrix. It can be shown that this power series is convergent for every $\mathbf{A}$, so that $e^{\mathbf{A}}$ makes sense. Given the matrix $\mathbf{A} = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}$, prove that

$$e^{\mathbf{A}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

18) It is not always possible to find an LU-factorization for a matrix. Show that there is no nonzero value of satisfy

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ l & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{pmatrix},$$

no matter what the values of l, u_{12} and u_{22} are.

Remark. *If $l_{ii} = 0$, then either the factorization is not possible, or if A is singular then there is an infinite number of possible factorization.* by writing $U\mathbf{x} = \mathbf{y}$, and then putting $L\mathbf{y} = b$ and solving for $\mathbf{y}$.

19) Determine whether the following subset of R^n is a subspace of R^n.

$$\{\mathbf{x} \text{ such that } \sum_{j=1}^{n} x_j = 0\}.$$

20) Consider the bases $S_1 = \{e_1, e_2, e_3\}$ and $S_2 = \{f_1, f_2, f_3\}$ for R^3, where $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$, and $f_1 = (1,1,1)$, $f_2 = (1,1,0)$, and $f_3 = (1,0,0)$. Find the transition matrix from S_1 to S_2.

21) Find the orthogonal projection of $\mathbf{b}$ onto $\mathcal{M} = span\{\mathbf{u}\}$, and then determine the orthogonal projection of $\mathbf{b}$ onto $\mathcal{M}^T$, where

$$\mathbf{b} = \begin{pmatrix} 4 \\ 8 \end{pmatrix} \qquad \text{and} \qquad \mathbf{u} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

22) Determine the orthogonal projection of $\mathbf{b}$ onto $\mathcal{M}$, where

$$\mathbf{b} = \begin{pmatrix} 5 \\ 2 \\ 5 \\ 3 \end{pmatrix} \qquad \text{and} \qquad \mathcal{M} = span\left\{ \begin{pmatrix} -3/5 \\ 0 \\ 4/5 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 4/5 \\ 0 \\ 3/5 \\ 0 \end{pmatrix} \right\}.$$

23) Describe all 2×2 orthogonal projections in $R^{2\times 2}$.

24) Find the **QR**-decomposition of $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

25) Determine the **QR** factors of $\begin{pmatrix} 0 & -20 & -14 \\ 3 & 27 & -4 \\ 4 & 11 & -2 \end{pmatrix}$.

26) Use MATLAB to find the least-squares solution of the equation $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 7 \\ 2 \\ 3 \\ 6 \\ 5 \\ 4 \end{pmatrix}.$$

References

[1] M. A. Akivis, V. V. Goldberg, *An Introduction to Linear Algebra and Tensors*, Dover, NY, 1977

[2] C. G. Cullen, Matrices and Linear Transformations, Dover, NY, 1990

[3] H. Eves, *Elementary Matrix Theory*, Dover, 1980

[4] J. Goldberg, Matrix Theory with Applications, McGraw-Hill, NY, 1991

[5] S. I. Grossman, *Elementary Linear Algebra*, Wadsworth Publishing Company, 1987

[6] I. N. Hersteen, I. Kaplansky, *Matters Mathematical*, Harper and Row Publishers, NY, 1974

[7] L. W. Johnson, R. D. Riess, J. T. Arnold, Introduction to Linear Algebra, Addison-Wesley, MA, 1989

[8] A. D. Krauss, *Matrices for Engineers*, Hemisphere Publishing Corporation, Washington, 1987

[9] E. Kreyszig, *Advanced Engineering Mathematics*, John Wiley, NY, 1964

[10] S. Leon, E. Herman, R. Faulkenberry, ATLAST, Computer Exercises for Linear Algebra, Prentice-Hall, NJ, 1996

[11] M. Lowengrub, J. G. Stampfli, *Topics in Calculus*, Xerox College Publishing, Lexengton, MA, 1975

[12] M. Marcus, A Survey of Finite Mathematics, Dover, NY, 1993

[13] M. Marcus, H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Dover, NY, 1992

[14] S. Leon, E. Herman, R. Faulkenberry, ATLAST, Computer Exercises for Linear Algebra, Prentice-Hall, NJ, 1996

[15] J. E. Maxfield, Keys to Mathematics, W. B. Saunders Campany, Philadelphia, 1973

[16] B. Roberts, D. L. Schulze, *Modern Mathematics and Economic Analysis*, W. W. Norton, Inc. NY, 1973

[17] H. Schneider, G. P. Baker, *Matrices and Linear Algebra*, Dover, NY, 1989

[18] G. E. Shilov, *Linear Algebra*, Dover, NY, 1977

[19] E. L. Spitznagel, Jr, *Selected Topics in Mathematics*, Holt, Rinehart and Winston, Inc. NY, 1971

[20] K. A. Stroud, Further Engineering Mathematics, Springer-Verlag, NY, 1986

[21] F. Zhang, *Linear Algebra– Challenging Problems for the Student*, The John Hopkins University Press, Baltimore, 1996