MATH 247, FALL 2004

Characterisation of Positive Definite Matricies

The goal of this is to supplement the Appendix in our Lecture Notes. We aim to give a proof of Theorem A.3.8, accessible to any student who has had a good course in linear algebra.

Recall from Appendix A.3:

Theorem A.3.6 *If A in $M_N(\mathbb{R})$ is symmetric, then the following are equivalent:*

(i) *all of the eigenvalues of A are positive, i.e. A is* positive definite.

(ii) $Ax \cdot x > 0$ *for every x in* $\mathbb{R}^N \setminus \{0\}$.

We recall that the *eigenvalues* are exactly the roots of the characteristic polynomial

$$p_A(t) = \det(A - tI).$$

The *multiplicity* of a root λ is the number of times the factor $(t - \lambda)$ divides $p_A(t)$.

We import the following from linear algebra which is a refomulation of Theorem A.3.4:

Diagonalisation Theorem *If A in $M_N(\mathbb{R})$ is symmetric, then:*

(i) *all of the eigenvalues of A are real, so we may order them $\lambda_1 \leq \cdots \leq \lambda_N$ (counting multiplicity), and*

(ii) *there is a matrix U in $M_N(\mathbb{R})$ which is* orthogonal, *so $U^t = U^{-1}$, and*

$$U^{-1}AU = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}.$$

We introduce a *partial order* in on symmetric matricies

$$A \prec B \quad \Leftrightarrow \quad B - A \ \text{ is positive definite}$$

It is called a "partial" order, since it may be possible, for a given pair of symmetric matrices, that neither $A \prec B$ nor $B \prec A$. Thus we obtain the following consequence of Theorem A.3.6.

Corollary A *If A, B in $M_N(\mathbb{R})$ are symmetric then $A \prec B$ if and only if $(B-A)x \cdot x > 0$ for each x in $\mathbb{R}^N$.*

The following is an exercise.

Corollary B *If $A, B \in M_N(\mathbb{R})$ and U in $M_N(\mathbb{R})$ is orthogonal, then $A \prec B$ if and only if $U^{-1}AU \prec U^{-1}BU$.*

We introduce the handy notation $\prec$ so that we may easily state the following.

Inversion Lemma *If A in $M_N(\mathbb{R})$ is symmetric and positive definite, and $\lambda > 0$, then*

$$(A + \lambda I)^{-1} \prec A^{-1}.$$

Proof. If U is the orthogonal matrix promised by the Diagonalisation Theorem then

$$U^{-1}AU = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_N \end{bmatrix} \quad \text{and} \quad U^{-1}(A+\lambda I)U = \begin{bmatrix} \lambda_1 + \lambda & \dots & 0 \\ \vdots & \ddots & 0 \\ 0 & \dots & \lambda_N + \lambda \end{bmatrix}.$$

Thus we have that

$$U^{-1}A^{-1}U = \left(U^{-1}AU\right)^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_N} \end{bmatrix} \quad \text{and} \quad U^{-1}(A+\lambda I)^{-1}U = \begin{bmatrix} \frac{1}{\lambda_1+\lambda} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_N \lambda} \end{bmatrix}.$$

Hence

$$U^{-1}A^{-1}U - U^{-1}(A+\lambda I)^{-1}U = \begin{bmatrix} \frac{1}{\lambda_1} - \frac{1}{\lambda_1+\lambda} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_N} - \frac{1}{\lambda_N+\lambda} \end{bmatrix}$$

is positive definite, whence $U^{-1}(A+\lambda I)^{-1}U \prec U^{-1}A^{-1}U$ so $(A+\lambda I)^{-1} \prec A^{-1}$, by Corollary B. $\square$

We now get to our main theorem, which allows us to detect whether a matrix is positive definite, or not.

Postive Definiteness Test *Let A in $M_N(\mathbb{R})$ be symmetric. Then the following are equivalent.*

(i) *A is positive definite.*

(ii) *For each $k = 1, 2, \ldots, N$, let A_k in $M_k(\mathbb{R})$ be the matrix A, truncated to the first k rows and first k columns. Then $\det A_k > 0$ for each $k = 1, \ldots, N$.*

Proof. (i)⇒(ii) By the Diagonalisation Theorem we see that

$$\det A = \det\bigl(U^{-1}AU\bigr) = \lambda_1\lambda_2 \ldots \lambda_N > 0$$

since all of the eigenvalues are positive. Thus we have the statement for $k = N$.

If $k < N$ observe that for each $x \in \mathbb{R}^k \setminus \{0\}$

$$A_k x \cdot x = A \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} > 0$$

so A_k is positive definite. Hence $\det A_k > 0$, as above.

(ii)⇒(i) We use induction on N. For $N = 1$, the result is obvious. Now suppose that $N > 1$, that A_{N-1} is positive definite and that $\det A > 0$.

Suppose that $\lambda \leq 0$. Using the proof of the Inversion Lemma it is easy to see that $A_{N-1} - \lambda I_{N-1}$ is positive definite, thus invertible, and hence $(A_{N-1} - \lambda I_{N-1})^{-1}$ is positive definite too. Moreover, the Inversion Lemma tells us that

$$(A_{N-1} - \lambda I_{N-1})^{-1} \preceq A_{N-1}^{-1} \preceq A_{N-1}^{-1} - \frac{\lambda}{x \cdot x} I_{N-1} \qquad (\clubsuit)$$

where $x \in \mathbb{R}^{N-1} \setminus \{0\}$. Now we write A as a block matrix

$$A = \begin{bmatrix} A_{N-1} & B^t \\ B & a_{NN} \end{bmatrix}$$

where $B = [a_{N1} \ \dots \ a_{N,N-1}]$, and we compute $p_A(\lambda)$:

$$\begin{aligned}
\det(A - \lambda I_N) &= \det \begin{bmatrix} A_{N-1} - \lambda I_{N-1} & B^t \\ B & a_{NN} - \lambda \end{bmatrix} \\
&= \det \left(\begin{bmatrix} I_{N-1} & 0 \\ -B(A_{N-1} - \lambda I_{N-1})^{-1} & 1 \end{bmatrix} \begin{bmatrix} A_{N-1} - \lambda I_{N-1} & B^t \\ B & a_{NN} - \lambda \end{bmatrix} \right) \\
&= \det \begin{bmatrix} A_{N-1} - \lambda I_{N-1} & B^t \\ 0 & -B(A_{N-1} - \lambda I_{N-1})^{-1} B^t + a_{NN} - \lambda \end{bmatrix} \\
&= \det(A_{N-1} - \lambda I_{N-1}) \left(-B(A_{N-1} - \lambda I_{N-1})^{-1} B^t + a_{NN} - \lambda \right).
\end{aligned}$$

[This clever matrix trick is the heart of the proof. The rest are "standard" details.] Setting $\lambda = 0$ we see that

$$0 < \det A = \det A_{N-1} \left(-B A_{N-1}^{-1} B^t + a_{NN} \right).$$

Since $\det A_{N-1} > 0$, it follows that

$$a_{NN} > B A_{N-1}^{-1} B^t \geq 0. \qquad (\spadesuit)$$

Now we suppose, again, that $\lambda \leq 0$. If $B = 0$ we see that $\det(A - \lambda I_N) > 0$, so A has only positive eignevalues, and hence must be positive definite. If $B \neq 0$ we set $x = (a_{N1}, \dots, a_{N,N-1})$ and we see that

$$B(A_{N-1} - \lambda I_{N-1})^{-1} B^t = (A_{N-1} - \lambda I_{N-1})^{-1} x \cdot x.$$

It follows from ($\clubsuit$) and Corollary A, above, that

$$(A_{N-1} - \lambda I_{N-1})^{-1} x \cdot x \leq \left(A_{N-1}^{-1} - \frac{\lambda}{x \cdot x} I_{N-1} \right) x \cdot x = A_{N-1}^{-1} x \cdot x - \lambda = B A_{N-1}^{-1} B^t - \lambda.$$

It then follows from ($\spadesuit$) that

$$\begin{aligned}
-B\big(A_{N-1} - \lambda I_{N-1}\big)^{-1} B^t + a_{NN} - \lambda &\geq -\big(B A_{N-1}^{-1} B^t - \lambda\big) + a_{NN} - \lambda \\
&= a_{NN} - B A_{N-1}^{-1} B^t > 0.
\end{aligned}$$

Thus we have that

$$\det(A - \lambda I_N) \geq \det(A_{N-1} - \lambda I_{N-1}) \left(a_{NN} - B A_{N-1}^{-1} B^t \right) > 0$$

and again we see that A can have only positive eigenvalues, and hence must be positive definite. □

I am grateful to Heydar Radjavi for showing me this proof.