10. Quadratic Forms

10.1 Definition: Let U be a vector space over a field $\mathbb{F}$. A **quadratic form** on U is a map $K : U \to \mathbb{F}$ of the form $K(u) = L(u,u)$ for some symmetric bilinear form L on U. Note that for $u, v \in U$ we have

$$K(u+v) = L(u+v,u+v) = L(u,u)+L(u,v)+L(v,u)+L(v,v) = K(u)+2L(u,v)+K(v)$$

and so when $\text{char}\,(\mathbb{F}) \neq 2$ we have the **Polarization Identity**

$$L(u,v) = \tfrac{1}{2}\big(K(u+v) - K(u) - K(v)\big).$$

This shows that L is uniquely determined from K. Given a basis $\mathcal{A}$ for U, we define the **matrix** of K with respect to $\mathcal{A}$ to be the matrix of its (unique) associated symmetric bilinear form L, that is

$$[K]_{\mathcal{A}} = [L]_{\mathcal{A}}$$

so that for $u \in U$ we have

$$K(u) = [u]_{\mathcal{A}}{}^T [K]_{\mathcal{A}} [u]_{\mathcal{A}}.$$

When $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$, the matrix $A = [K]_{\mathcal{A}} \in M_n(\mathbb{F})$ has entries $A_{i,j} = L(u_j, u_i)$, and writing $x = [u]_{\mathcal{A}}$ we have

$$K(u) = x^T A x = \sum_{i=1}^n x_i A_{i,j} x_j = \sum_{i=1}^n A_{i,i} {x_i}^2 + 2 \sum_{i<j} A_{i,j} x_i x_j .$$

When we diagonalize the symmetric matrix A by choosing an invertible matrix $P \in M_n(\mathbb{F})$ so that $P^T A P = D = \text{diag}(d_1, d_2, \cdots, d_n)$, if we write $x = Pt$, or equivalently $t = P^{-1}x$, then we have

$$K(u) = x^T A x = t^T P^T A P\, t = t^T D t = \sum_{i=1}^n d_i {t_i}^2 .$$

10.2 Example: A polynomial or power series in the variables $x_1, x_2, \cdots, x_n$ can be written as

$$f(x) = \sum_{i_1, i_2, \cdots, i_n \geq 0} c_{i_1, \cdots, i_n} {x_1}^{i_1} {x_2}^{i_2} \cdots {x_n}^{i_n} = \sum_{k \geq 0} f_k(x)$$

where

$$f_k(x) = \sum_{\substack{i_1, i_2, \cdots, i_n \geq 0 \\ i_1 + i_2 + \cdots + i_n = k}} c_{i_1, \cdots, i_n} {x_1}^{i_1} \cdots {x_n}^{i_n} .$$

For each $k \geq 0$, the polynomial $f_k(x)$ is a **homogeneous** polynomial of degree k, which means that $f_k(tx) = t^k f_k(x)$ for all $x \in \mathbb{R}^n$ and all $t \in \mathbb{R}$. By relabeling the coefficients, we can also write

$$f_0(x) = a\ ,\ f_1(x) = \sum_{1 \leq i \leq n}^{n} a_i x_i\ ,\ f_2(x) = \sum_{1 \leq i \leq j \leq n} a_{i,j}\, x_i x_j\ ,\ f_3(x) = \sum_{1 \leq i \leq j \leq k \leq n} a_{i,j,k}\, x_i x_j x_k$$

and so on. In particular, when $\text{char}\,(\mathbb{F}) \neq 2$, we have $f_2(x) = \sum\limits_{1 \leq i \leq j \leq n} a_{i,j}\, x_i x_j = x^T A x$ where $A \in M_n(\mathbb{F})$ is the matrix with entries $A_{i,i} = a_i$ for $1 \leq i \leq n$ and $A_{i,j} = A_{j,i} = \frac{1}{2} a_{i,j}$ for $1 \leq i < j \leq n$. Thus a quadratic form on $\mathbb{F}^n$ is the same thing as a homogeneous polynomial in $x_1, \cdots, x_n$ of degree 2.

10.3 Example: Sketch or describe the curve $3x^2 - 4xy + 6y^2 = 10$.

Solution: Let K be the quadratic form on $\mathbb{R}^2$ given by $K(x,y) = 3x^2 - 4xy + 6y^2$. Note that

$$K(x,y) = (x\ y)A\begin{pmatrix}x\\y\end{pmatrix} \quad \text{where} \quad A = \begin{pmatrix}3 & -2\\-2 & 6\end{pmatrix}.$$

The characteristic polynomial of A is

$$f_A(x) = \det\big(A - xI\big) = \det\begin{pmatrix}3-x & -2\\-2 & 6-x\end{pmatrix} = (x^2 - 9x + 14 = (x-7)(x-2)$$

so the eigenvalues of A are $\lambda_1 = 7$ and $\lambda_2 = 2$. We have

$$A - \lambda_1 I = \begin{pmatrix}-4 & -2\\-2 & -1\end{pmatrix} \cong \begin{pmatrix}2 & 1\\0 & 0\end{pmatrix}$$

and so $u_1 = \frac{1}{\sqrt{5}}(1,-2)^T$ is a unit eigenvector for λ_1. Since A is symmetric, the eigenspace of λ_2 is orthogonal to the eigenspace of λ_1, so $u_2 = \frac{1}{\sqrt{5}}(2,1)^T$ is a unit eigenvector for λ_2. Thus we have

$$P^TAP = D \;\text{ where }\; P = \big(u_1, u_2\big) = \tfrac{1}{\sqrt{5}}\begin{pmatrix}1 & 2\\-2 & 1\end{pmatrix} \;\text{ and }\; D = \begin{pmatrix}\lambda_1 & 0\\0 & \lambda_2\end{pmatrix} = \begin{pmatrix}7 & 0\\0 & 2\end{pmatrix}.$$

We make a change in coordinates by writing $\begin{pmatrix}x\\y\end{pmatrix} = P\begin{pmatrix}s\\t\end{pmatrix}$, and then we have

$$K(x,y) = (x\ y)A\begin{pmatrix}x\\y\end{pmatrix} = (s\ t)P^TAP\begin{pmatrix}s\\t\end{pmatrix} = (s\ t)D\begin{pmatrix}s\\t\end{pmatrix} = 7s^2 + 2t^2$$

and so

$$K(x,y) = 10 \iff 7s^2 + 2t^2 = 10 \iff \frac{s^2}{10/7} + \frac{t^2}{5} = 1.$$

Thus, in the st-plane, the curve is the ellipse with vertices at $\pm\big(\sqrt{\frac{10}{7}}, 0\big)$ and $\pm\big(0, \pm\sqrt{5}\big)$. Since our change of coordinate matrix P is an orthogonal matrix it preserves inner product, norm and angle (indeed P is a rotation matrix), in the xy-plane, the curve is an ellipse of the same shape with vertices at

$$P\begin{pmatrix}\pm\sqrt{\frac{10}{7}}\\0\end{pmatrix} = \pm\sqrt{\tfrac{10}{7}}\cdot\tfrac{1}{\sqrt{5}}\begin{pmatrix}1 & 2\\-2 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \pm\sqrt{\tfrac{2}{7}}\begin{pmatrix}1\\-2\end{pmatrix}\ , \text{ and}$$
$$P\begin{pmatrix}0\\\pm\sqrt{5}\end{pmatrix} = \pm\sqrt{5}\cdot\tfrac{1}{\sqrt{5}}\begin{pmatrix}1 & 2\\-2 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \pm\begin{pmatrix}2\\1\end{pmatrix}.$$

10.4 Theorem: *Let U be an n-dimensional inner product space over $\mathbb{R}$ and let $K : U \to \mathbb{R}$ be a quadratic form on U. Let $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ be an orthonormal basis for U such that $[K]_{\mathcal{A}} = D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then*

$$\max_{u \in U, |u|=1} K(u) = K(u_1) = \lambda_1 \quad \text{and} \quad \min_{u \in U, |u|=1} K(u) = K(u_n) = \lambda_n.$$

Proof: Let $u \in U$ and write $x = [u]_{\mathcal{A}}$. Note that $|x| = |u|$ since $\mathcal{A}$ is orthonormal. When $|u| = |x| = 1$ we have

$$K(u) = x^T D x = \sum_{i=1}^{n} \lambda_i {x_i}^2 \leq \sum_{i=1}^{n} \lambda_1 {x_i}^2 = \lambda_1 \sum_{i=1}^{n} {x_i}^2 = \lambda_1 |x|^2 = \lambda_1$$

and when $u = u_1$ we have $x = [u_1]_{\mathcal{A}} = e_1$ so that $K(u) = K(u_1) = {e_1}^T D e_1 = \lambda_1$. This shows that $\max_{u \in U, |u|=1} K(u) = K(u_1) = \lambda_1$, and the proof that $\min_{u \in U, |u|=1} K(u) = K(u_n) = \lambda_n$ is similar.

10.5 Theorem: *Let U and V be finite dimensional inner product spaces over $\mathbb{R}$ and let $L : U \to V$ be a linear map. Then*

$$\max_{u \in U, |u|=1} \big|L(u)\big| = \big|L(u_1)\big| = \sigma_1 \quad \text{and} \quad \min_{u \in U, |u|=1} \big|L(u)\big| = \big|L(u_n)\big| = \sigma_n$$

where σ_1 and σ_n are the largest and smallest singular values of L and u_1 and u_n are unit eigenvectors of the map L^*L for the eigenvalues $\lambda_1 = {\sigma_1}^2$ and $\lambda_n = {\sigma_n}^2$.

Proof: Choose an orthonormal basis $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ for U such that

$$[L^*L]_{\mathcal{A}} = D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of L^*L with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and let $\sigma_i = \sqrt{\lambda_i}$. Choose any orthonormal basis $\mathcal{B}$ for V and let $A = [L]_{\mathcal{B}}^{\mathcal{A}}$. Note that

$$A^T A = A^* A = [L^*]_{\mathcal{A}}^{\mathcal{B}} [L]_{\mathcal{B}}^{\mathcal{A}} = [L^*L]_{\mathcal{A}} = D.$$

Let $u \in U$ and write $x = [u]_{\mathcal{A}}$. Note that $|x| = |u|$ and

$$\big|L(u)\big|^2 = \big|[L(u)]_{\mathcal{B}}\big|^2 = \big|[L]_{\mathcal{B}}^{\mathcal{A}}[u]_{\mathcal{A}}\big|^2 = \big|Ax\big|^2 = (Ax)^T(Ax) = x^T A^T A\, x = x^T D x.$$

As in the proof of the previous theorem, we see that

$$\max_{u \in U, |u|=1} \big|L(u)\big|^2 = \big|L(u_1)\big|^2 = \lambda_1 \quad \text{and} \quad \min_{u \in U, |u|=1} \big|L(u)\big|^2 = \big|L(u_n)\big|^2 = \lambda_n$$

and so

$$\max_{u \in U, |u|=1} \big|L(u)\big| = \big|L(u_1)\big| = \sigma_1 \quad \text{and} \quad \min_{u \in U, |u|=1} \big|L(u)\big| = \big|L(u_n)\big| = \sigma_n.$$

10.6 Theorem: *Let U and V be non-trivial subspaces of $\mathbb{R}^n$ with $U \cap V = \{0\}$. Then $\theta(U, V) = \cos^{-1}(\sigma_1)$ where σ_1 is the largest singular value of the map $P : U \to V$ given by $P(x) = \mathrm{Proj}_V(x)$.*

Proof: The proof is left as an exercise.

10.7 Example: Let $U \subseteq \mathbb{R}^n$ be an open set with $a \in U$. Let $f : U \to \mathbb{R}$ be smooth (meaning that the partial derivatives of all orders all exist in U). The Taylor polynomial of degree 2 centred at $x = a$ for a smooth function of the variables $x_1, x_2, \cdots, x_n$ can be written as

$$T(x) = f(a) + Df(a)\,(x-a) + (x-a)^T Hf(a)\,(x-a)$$

where $Df(a) \in M_{1\times n}(\mathbb{R})$ and $Hf(a) \in M_{n\times n}(\mathbb{R})$ are the matrices with entries

$$Df(a)_i = \frac{\partial f}{\partial x_i}(a) \quad \text{and} \quad Hf(a)_{i,j} = \frac{\partial^2 f}{\partial x_i\,\partial x_j}(a).$$

10.8 Theorem: (*The Second Derivative Test*) *Let $U \subseteq \mathbb{R}^n$ be an open set with $a \in U$. Let $f : U \to \mathbb{R}$ be smooth. Suppose that $Df(a) = 0$. Then*

(1) if $Hf(a)$ is positive definite then $f(x)$ has a local minimum at $x = a$,
(2) if $Hf(a)$ is negative definite then $f(x)$ has a local maximum at $x = a$, and
(3) if $Hf(a)$ is indefinite then $f(x)$ has a saddle point at $x = a$.

Proof: We omit the proof. This theorem is often proven in a calculus course.

10.9 Remark: Let U, V and W be a vector spaces over $\mathbb{C}$. A map $L : U \times V \to W$ is called **sesquilinear** when

$$\begin{array}{lcl} L(x_1+x_2,y) = L(x_1,y) + L(x_2,y) & , & L(tx,y) = t\,L(x,y)\,, \\ L(x,y_1+y_2) = L(x,y_1) + L(x,y_2) & \text{and} & L(x,ty) = \overline{t}\,L(x,y) \end{array}$$

for all $x, x_1, x_2 \in U$, and all $y, y_1, y_2 \in V$ and for all $t \in \mathbb{C}$. A **Hermitian form** is a sesquilinear map $L : U \times U \to \mathbb{C}$ with the property that $L(y,x) = \overline{L(x,y)}$ for all $x, y \in U$, and a **skew-Hermitian form** is a sesquilinear map $L : U \times U \to \mathbb{C}$ such that $L(y,x) = -\overline{L(x,y)}$ for all $x, y \in U$.

10.10 Example: As an exercise, think about how the theory of bilinear and quadratic forms, from this and the previous chapter, carry over to Hermitian forms.