16 Positive (semi)definite symmetric matrices

Recall (from Lecture 6) that a function $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is an *inner product* if it satisfies the following properties:

- B is a *bilinear form*:
  - $B(u + v, w) = B(u, w) + B(v, w)$ for $u, v, w \in \mathbb{R}^n$;
  - $B(u, v + w) = B(u, v) + B(u, w)$ for $u, v, w \in \mathbb{R}^n$;
  - $B(\lambda u, v) = \lambda B(u, v)$ for $\lambda \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$;
  - $B(u, \lambda v) = \lambda B(u, v)$ for $\lambda \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$.
- B is *symmetric*: $B(u, v) = B(v, u)$ for all u, v in $\mathbb{R}^n$.
- B is *positive definite*: $B(u, u) > 0$ for all non-zero $u \in \mathbb{R}^n$.

All of this taken together says that an *inner product* on a real vector space V is a *positive definite symmetric bilinear form*. For this definition there is no requirement that V be finite dimensional although that will be the case of interest to us.

Theorem 16.1. *If* B *is an inner product on* V*, then the function* $||.||_B : V \to \mathbb{R}$ *defined by* $||v|| = \sqrt{B(v, v)}$ *is a vector norm.*

Proof coming later.

Now suppose that B is a bilinear form (as defined above) on a vector space of dimension n over a field $\mathbb{F}$, and let $\mathcal{B} = \{b_1, \dots, b_n\}$ be a basis of V. Let $u = \sum u_i b_i$ and $v = \sum v_i b_i$ be vectors in V. (We write u and v also for the column coordinate vectors with entries $u_1, \dots, u_n$ and $v_1, \dots, v_n$). Write a_{ij} for the real number $B(b_i, b_j)$, and let A be the $n \times n$ matrix whose (i, j)-entry is a_{ij}. Then

$$\begin{aligned}
B(u, v) &= B(u_1 b_1 + \cdots + u_n b_n, v) \\
&= \sum_{i=1}^{n} B(u_i b_i, v) \\
&= \sum_{i=1}^{n} u_i B(b_i, v) \\
&= \sum_{i=1}^{n} u_i B(b_i, v_1 b_1 + \cdots + v_n b_n) \\
&= \sum_{i=1}^{n} u_i \sum_{j=1}^{n} B(b_i, v_j b_j) \\
&= \sum_{i=1}^{n} u_i \sum_{j=1}^{n} v_j B(b_i, b_j) \\
&= \sum_{i=1}^{n} u_i \sum_{j=1}^{n} a_{ij} v_j \\
&= \sum_{i=1}^{n} u_i (Av)_i \\
&= u^T A v.
\end{aligned}$$

Definition 16.2. *The matrix* A *is called the* Gram matrix *of the bilinear form* B *with respect to the basis* $\{b_1, \dots, b_n\}$*. Its* (i, j) *entry is* $B(b_i, b_j)$*. For any pair of vectors* u *and* v*, we have*

$$B(u, v) = [u]^T A [v],$$

where $[u]$ *and* $[v]$ *respectively denote the column coordinate vectors of* u *and* v *with respect to* $\{b_1, \dots, b_n\}$.

Now suppose that $\mathcal{B}$ and $\mathcal{C}$ are different bases of the vector space on which B is defined, and let P be the change of basis matrix from $\mathcal{C}$ to $\mathcal{B}$, so $[v]_{\mathcal{B}} = P[v]_{\mathcal{C}}$ for any vector v. Let $A_{\mathcal{B}}$ and $A_{\mathcal{C}}$ be the Gram matrices of B with respect to $\mathcal{B}$ and $\mathcal{C}$ respectively. The relationship (in matrix algebra terms) between these two matrices can be identified as follows. Let u and v be vectors. Then

$$\begin{aligned} B(u,v) &= [u]_{\mathcal{B}}^T A_{\mathcal{B}} [v]_{\mathcal{B}} \\ &= (P[u]_{\mathcal{C}})^T A_{\mathcal{B}} P[u]_{\mathcal{C}} \\ &= [u]_{\mathcal{C}}^T P^T A_{\mathcal{B}} P[u]_{\mathcal{C}}. \end{aligned}$$

We conclude that $A_{\mathcal{C}} = P^T A_{\mathcal{B}} P$, where P is the change of basis matrix from $\mathcal{C}$ to $\mathcal{B}$. This motivates the following definition.

Definition 16.3. *Let* A *and* B *be square matrices. Then* A *and* B *are* congruent *if there exists an invertible matrix* P *for which*

$$B = P^T A P.$$

Note that congruence is an equivalence relation on $M_n(\mathbb{F})$ for a field $\mathbb{F}$. It has the role for bilinear forms that similarity has for linear transformations. If two matrices are congruent they are Gram matrices for the same bilinear form with respect to different bases. If two matrices are similar they represent the same linear transformation with respect to different bases. Given a basis, a matrix may be considered to represent a linear transformation or a bilinear form. In the first case, the equivalence relation of interest is similarity, in the second case it is congruence. It is possible for two matrices to be congruent but not similar, or similar but not congruent, or neither, or both. Note that if P is an orthogonal matrix (so $P^T = P^{-1}$) then the expression $B = P^T A P$ simultaneously demonstrates the similarity and the congruence of A and B. In particular we have shown that every real symmetric matrix is both similar and congruent to the same diagonal matrix.