THE GRAM MATRIX, ORTHOGONAL PROJECTION, AND VOLUME

1. The Gram Matrix

Given a list of vectors, $u_1, ..., u_k \in \mathbb{R}^n$, the matrix

$$\begin{pmatrix} u_1 \cdot u_1 & \dots & u_1 \cdot u_k \\ \vdots & \dots & \vdots \\ u_k \cdot u_1 & \dots & u_k \cdot u_k \end{pmatrix} \tag{1.1}$$

is called the Gram matrix. We will denote it $gm(u_1, ..., u_k)$. The Gram determinant is the determinant of the Gram matrix, denoted $gd(u_1, ..., u_k)$. If we let Y denote the $n \times k$ matrix with columns $u_1, ..., u_k$, then

$$gm(u_1, ..., u_k) = Y^T Y \tag{1.2}$$

$$gd(u_1, ..., u_k) = det(Y^T Y)\ . \tag{1.3}$$

Proposition 1. *The vectors $u_1, ..., u_k$ are linearly independent if and only if their Gram matrix is nonsingular.*

Proof. Assume $u_1, ..., u_k$ are linearly dependent. Let $\mathbf{a} = (a_1, ..., a_k)$ be a nonzero vector such that $\sum_{i=1}^k a_i u_i = 0$. Then for all j,

$$\sum_{i=1}^k a_i u_j \cdot u_i = 0\ , \tag{1.4}$$

so $\mathbf{a}^T$ is in the kernel of the Gram matrix.

Conversely, if the Gram matrix is singular, then there exists a nonzero vector $\mathbf{a} = (a_1, ..., a_k)$ such that (1.4) holds. Let $\mathbf{w} = \sum_{i=1}^k a_i u_i$. Then $\mathbf{w}$ is orthogonal to every u_j, and therefore orthogonal to itself. That is, $\mathbf{w} \cdot \mathbf{w} = 0$. Then $\mathbf{w} = \mathbf{0}$. Thus, $u_1, ..., u_k$ are linearly dependent. □

2. Orthogonal projection

Let $W \subset \mathbb{R}^n$ be a subspace. The *orthogonal complement* of W is the subspace

$$W^\perp = \{v \in \mathbb{R}^n \mid \forall w \in W, v \cdot w = 0.\} \tag{2.1}$$

Proposition 2. $\mathbb{R}^n = W \oplus W^\perp$.

Proof. One must show that for all $v \in \mathbb{R}^n$ there exists unique $w \in W$ such that $v - w \in W^\perp$. Let $u_1, ..., u_k$ be a basis for W. The desired vector w is of the form $(u_1, ..., u_k)\mathbf{a}$, for some $\mathbf{a} \in \mathbb{R}^k$. Then $v-w \in W^\perp$ if and only if, for all j,

$$u_j \cdot (v - \sum a_i u_i) = 0 \ ,$$

i.e,

$$gm(u_1, ..., u_k)\mathbf{a} = (u_1 \cdot v, ..., u_k \cdot v)^T \ .$$

The vectors $u_1, ..., u_k$ are linearly independent, so by the lemma, this equation has a unique solution.

□

3. Volume

The vectors $u_1, ..., u_k$ span a parallelopiped,

$$P(u_1, ..., u_k) = \{\sum a_i u_i \mid \forall i, 0 \le a_i \le 1 \ .\}$$

A one-dimensional parallelopiped is a line segment and a two-dimensional parallelopiped is a parallelogram. An obvious way to define the volume of the parallelopiped is by recursion:

$$Vol(u) = \sqrt{u \cdot u} \tag{3.1}$$

$$Vol(u_1..., u_k, u_{k+1}) = Vol(u_1, ..., u_k)||\tilde{u}_{k+1}|| \ , \tag{3.2}$$

where $\tilde{u}_{k+1}$ is the orthogonal projection of u_{k+1} onto $span(u_1, .., u_k)^\perp$. In other words,

$$\text{Volume} = (\text{Volume of base}) \times \text{height} \ .$$

This is of course the natural extension of the notion of volume in three dimensions. It is not immediately evident from the definition that the volume function is independent of the ordering of the vectors, nor is it obvious that volume is invariant under rotation. The next theorem shows that both of these intuitive properties indeed hold.

Theorem 3.

$$Vol(u_1, ..., u_k)^2 = gd(u_1, ..., u_k) \ . \tag{3.3}$$

Proof. If $k = 1$, formula (3.3) reduces to (3.1). Assume then that formula (3.3) holds for k, and consider $u_1, ..., u_{k+1} \in \mathbb{R}^n$. Regard $(u_1, ..., u_k)$ as an $n \times k$ matrix Y. Then there exists a vector $c \in \mathbb{R}^k$ such that $u_{k+1} = Yc + \tilde{u}_{k+1}$. The matrix

$$\begin{pmatrix} Y^TY & Y^TYc \\ c^TY^TY & c^TY^TYc \end{pmatrix}$$

has vanishing determinant, because its last column is c times the first k columns. Furthermore,

$$gm(u_1, ..., u_{k+1}) = \begin{pmatrix} Y^TY & Y^TYc \\ c^TY^TY & c^TY^TYc + ||\tilde{u}_{k+1}||^2 \end{pmatrix}$$

Thus

$$gd(u_1, ..., u_{k+1}) = det \begin{pmatrix} Y^TY & Y^TYc \\ c^tY^TY & c^TY^TYc \end{pmatrix} + det(\begin{pmatrix} Y^TY & 0 \\ c^TY^TY & ||\tilde{u}_{k+1}||^2 \end{pmatrix} = gd(u_1, ..., u_k)||\tilde{u}_{k+1}||^2 \ .$$

□

Corollary 4.

1. Let $u_1, ..., u_k$ be vectors in $\mathbb{R}^k$. Then

$$Vol(u_1, ..., u_k) = |det(u_1, ..., u_k)| \ .$$

2. Let $u_1, ..., u_k$ be vectors in $\mathbb{R}^n$, and let A be an $n \times n$ orthogonal matrix. Then

$$Vol(Au_1, ..., Au_k) = Vol(u_1, ..., u_k) \ .$$

3. $Vol(u_1, ..., u_k)$ is a symmetric function of $u_1, ..., u_k$.

Proof. If Y is a square matrix, then $det(Y^TY) = det(Y^T)det(Y) = det(Y)^2$. This proves 1.

If A is an orthogonal matrix, then $A^TA = I$ by definition. Then $Y^TY = (AY)^T(AY)$, which proves 2.

To prove 3, let $Y = (u_1, ..., u_k)$ and let $(\sigma_1, ..., \sigma_k) \in S_k$ be a permutation. Then there is a $k \times k$ permutation matrix P such that $(u_{\sigma_1}, ..., u_{\sigma_k}) = YP$. Then

$$gd(u_{\sigma_1}, ..., u_{\sigma_k}) = det(P^TY^TYP) = det(Y^TYPP^T) = gd(u_1, ..., u_k) \ .$$

□

3.1. **An analogue of the Pythagorean Theorem.** Recall that if u and v are vectors in $\mathbb{R}^3$, then $Vol(u, v) = ||u \times v||$. This is an alternative to formula (3.3). The validity of both formulas is then a polynomial identity in the entries of a 3×2 matrix. If

$$Y = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} \ ,$$

the identity is

$$det(Y^TY) = (a^2+b^2+c^2)(d^2+e^2+f^2) - (ad+be+cf)^2 = (bf-ce)^2 + (cd-af)^2 + (ae-bd)^2 \ ,$$

which one can readily check. Thus the area of a pair of vectors in $\mathbb{R}^3$ turns out to be the length of a vector constructed from the three 2×2 minors of Y. A similar phenomenon occurs for an arbitrary list of vectors $u_1, ..., u_k$ in $\mathbb{R}^n$.

Theorem 5. *Let $u_1, ..., u_k$ be a list of vectors in $\mathbb{R}^n$, and let Y denote the $n \times k$ matrix with columns $u_1, ..., u_k$. Let $v \in \mathbb{R}^{\binom{n}{k}}$ denote a vector whose components are all the $k \times k$ minors of Y, (listed in arbitrary order.) Then*

$$Vol(u_1, ..., u_k) = ||v|| \ .$$

Proof. The theorem asserts the identity

$$det(Y^T Y) = \sum_I det(Y_I)^2 \ , \tag{3.4}$$

where the sum runs over all increasing multiindices

$$I = (1 \le i_1 < i_2 < \cdots < i_k \le n) \ .$$

Fix an $n \times k$ matrix Y. For all $n \times k$ matrices X, define

$$F(X) = det(Y^T X) \ .$$

Then F is an alternating multilinear function of the columns of X, and therefore admits an expansion of the form

$$F(X) = \sum_I c_I det(X_I) \ .$$

If we let $E_{(I)}$ denote the matrix with columns $e_{i_1}, \ldots, e_{i_k}$, then $c_I = F(E_{(I)})$. Furthermore

$$Y^T E_{(I)} = Y_I^T \ .$$

Thus

$$\begin{aligned} & det(Y^T X) = F(X) \\ & = \sum_I F(E_{(I)}) det(X_I) = \sum_I det(Y_I^T) det(X_I) \\ & = \sum_I det(Y_I) det(X_I) \ . \end{aligned} \tag{3.5}$$

The result follows by letting $X = Y$.

□