Linear Algebra Notes

1 The Gram Matrix

Definition 1.1. Let $v_1, v_2, \ldots, v_n$ be vectors in $\mathbb{R}^d$. Define the associated $n \times n$ **Gram Matrix**, G, by

$$G_{i,j} = \langle v_i, v_j \rangle.$$

Remark 1.1. If we let V be the matrix whose columns are $v_1, v_2, \ldots, v_n$, then we can write $G = V^t V$. This will come in handy when proving things about the Gram matrix.

Lemma 1.1. *The Gram matrix is symmetric and positive semi-definite.*

Proof. The symmetry of G follows from the symmetry of the inner product. Alternatively,

$$G^t = (V^t V)^t = V^t (V^{tt}) = V^t V = G.$$

Let x be any vector in $\mathbb{R}^n$. We then have

$$x^t G x = x^t V^t V x = \langle Vx, Vx \rangle = \|Vx\|^2 \geq 0,$$

so G is positive semi-definite. □

Lemma 1.2. *The rank of the Gram matrix is the dimension of the space spanned by $v_1, v_2, \ldots, v_n$ in $\mathbb{R}^d$.*

Proof. Let $x \in \mathbb{R}^n$ and suppose $Vx = 0$. Then $Gx = V^t V x = 0$ as well, so $\ker V \subseteq \ker G$. On the other hand, suppose $Gx = 0$. Multiplying on the left by x^t gives

$$x^t G x = 0 \iff x^t V^t V x = 0 \iff \|Vx\|^2 = 0 \iff Vx = 0,$$

so $\ker G \subseteq \ker V$. Since G and V have the same kernel, they also have the same rank by the rank-nullity theorem. □

2 The Rayleigh Quotient and the Min-Max Theorem

Definition 2.1. Let M be a symmetric $n \times n$ matrix and let x be any nonzero vector in $\mathbb{R}^n$. The **Rayleigh quotient**, $R(M, x)$ is defined by

$$R(M, x) = \frac{\langle x, Mx \rangle}{\|x\|^2}.$$

Lemma 2.1. *Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of M, repeated according to multiplicity. For any nonzero x we have*

$$R(M, x) \in [\lambda_1, \lambda_n].$$

The extreme values are obtained on the corresponding eigenvectors of M.

Proof. Since M is symmetric, there is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors $v_1, \ldots, v_n$ corresponding to the eigenvalues $\lambda_1 \le \cdots \le \lambda_n$. Write x in this basis as $x = \sum_{i=1}^n \xi_i v_i$. It's easy to see that $R(M, cx) = R(M, x)$ for any nonzero constant c, so we can take x to have unit norm for convenience, $\sum \xi_i^2 = 1$. The Rayleigh quotient is then given by

$$R(M, x) = \frac{\sum_{i=1}^n \lambda_i \xi_i^2}{\sum_{j=1}^n \xi_j^2} = \sum_{i=1}^n \lambda_i \xi_i^2.$$

From here it's clear that $R(M, x)$ is minimized when $\xi_i^2 = \delta_{1,i}$ and maximized when $\xi_i^2 = \delta_{n,i}$ and that these bounds are realized when x is the appropriate eigenvector. $\square$

Remark 2.1. Since the eigenvectors of M are mutually orthogonal, we can use the Rayleigh quotient to order the eigenvalues:

$$\lambda_i = \inf_{x \perp v_j,\ j<i} R(M, x).$$

Theorem 2.2. *Let M be a symmetric $n \times n$ with eigenvalues $\lambda_1 \le \cdots \le \lambda_n$. The eigenvalues are given by the expression*

$$\lambda_k = \min_{U\colon\ \dim U = k} \max_{x \in U \setminus \{0\}} R(M, x) = \max_{U\colon\ \dim U = n-k+1} \min_{x \in U \setminus \{0\}} R(M, x).$$

Proof. Let $v_1, \ldots, v_n$ be the eigenvectors associated to the (ordered) eigenvalues of M. For any k, the space spanned by $u_k, \ldots, u_n$ has dimension $n - k + 1$. If U is a subspace of dimension k, then these subspaces must have nontrivial intersection. There is then some nonzero vector $v = \sum_{i=k}^n c_i v_i$ in this intersection whose Rayleigh quotient is given by

$$R(M, v) = \frac{\sum_{i=k}^n \lambda_i c_i^2}{\sum_{i=k}^n c_i^2} \ge \lambda_k.$$

This holds for all v in this intersection, so for any U of dimension k we have

$$\max_{v \in U \setminus \{0\}} R(M, v) \ge \lambda_k.$$

Note that this maximum is attained since $R(M, v)$ is continuous in v and its values are determined by those v with norm 1, which form a compact set. Since this is true for all U of dimension k, we can take the infimum over all such U.

$$\inf_{\dim U = k} \max_{v \in U \setminus \{0\}} R(M, v) \ge \lambda_k.$$

Consider the space $U = \operatorname{span}\{v_1, \ldots, v_k\}$. For any $v = \sum_{i=1}^k c_i v_i$ in here we have

$$R(M, v) = \frac{\sum_{i=1}^k \lambda_i c_i^2}{\sum_{i=1}^k c_i^2} \le \lambda_k.$$

In particular, this inequality is saturated when $v = v_k$. The infimum is then attained and we have the equality

$$\min_{\dim U = k} \max_{v \in U \setminus \{0\}} R(M, v) = \lambda_k.$$

The same idea shows the max-min equality. The vectors $v_1, \dots, v_k$ span a space of dimension k, so any subspace U with dimension $n-k+1$ must intersect it nontrivially. Any $v = \sum_{i=1}^k c_i v_i$ in this intersection satisfies

$$R(M, v) = \frac{\sum_{i=1}^k \lambda_i c_i^2}{\sum_{i=1}^k c_i^2} \le \lambda_k \implies \min_{v \in U \setminus \{0\}} R(M, v) \le \lambda_k.$$

In particular, when $U = \text{span}\{v_k, \dots, v_n\}$ we have equality. We can then take the maximum over all U with dimension $n-k+1$ to obtain

$$\max_{\dim U = n-k+1} \min_{v \in U \setminus \{0\}} R(M, v) = \lambda_k.$$

□

3 The Cauchy Interlacing Theorem

Theorem 3.1. *Suppose A is an $n \times n$ symmetric matrix with eigenvalues $\lambda_1 \le \dots \le \lambda_n$. Let B be an $m \times m$ principal submatrix of A, i.e. a matrix obtained from A by deleting its i-th row and i-th column for some collection of i's. If B has eigenvalues $\beta_1 \le \dots \le \beta_k$ then*

$$\lambda_k \le \beta_k \le \lambda_{n+k-m}.$$

In particular, if $m = n-1$, then

$$\lambda_1 \le \beta_1 \le \lambda_2 \le \beta_2 \le \dots \le \beta_{n-1} \le \lambda_n.$$

Proof. Without loss of generality, we can rearrange the rows and columns of A so that

$$A = \begin{bmatrix} B & X^t \\ X & Z \end{bmatrix}$$

for some $(n-m) \times (n-m)$ matrices X and Z. Let $u_1, \dots, u_n$ be the (ordered) eigenvectors of A and let $v_1, \dots, v_m$ be the (ordered) eigenvectors of B. For any $1 \le k \le m$, define the subspaces

$$U = \text{span}\{u_k, \dots, u_n\}, \quad V = \text{span}\{v_1, \dots, v_k\}, \quad \widetilde{V} = \left\{ \begin{pmatrix} v \\ 0 \end{pmatrix} \in \mathbb{R}^n : v \in V \right\}.$$

The space U has dimension $n-k+1$ and $\widetilde{V}$ has dimension k, so these spaces must intersect nontrivially. There is then some $\tilde{v} \in U \cap \tilde{V}$ corresponding to some $v \in V$. This v satisfies

$$\tilde{v}^t A \tilde{v} = \begin{bmatrix} v & 0 \end{bmatrix} \begin{bmatrix} B & X^t \\ X & Z \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = v^t B v.$$

As in our proof of the min-max theorem, $\lambda_k = \min_{x \in U} \frac{x^t A x}{x^t x}$ and $\beta_k = \max_{x \in V} \frac{x^t B x}{x^t x}$. This gives

$$\lambda_k \le \frac{\tilde{v}^t A \tilde{v}}{\tilde{v}^t \tilde{v}} = \frac{v^t B v}{v^t v} \le \beta_k.$$

We use the same idea for the other inequality. Define the spaces

$$U = \operatorname{span}\{u_1, \ldots, u_{n+k-m}\}, \quad V = \operatorname{span}\{v_k, \ldots, v_m\}, \quad \tilde{V} = \left\{ \begin{pmatrix} v \\ 0 \end{pmatrix} \in \mathbb{R}^n : v \in V \right\}.$$

The space U has dimension $n+k-m$ and $\tilde{V}$ has dimension $m-k+1$, so they must intersect nontrivially. That is, there is some $\tilde{v} \in \tilde{V} \cap U$ corresponding to some $v \in V$. We again have by the min-max theorem

$$\lambda_{n+k-m} = \max_{x \in U} \frac{x^t A x}{x^t x} \geq \frac{\tilde{v}^t A \tilde{v}}{\tilde{v}^t \tilde{v}} = \frac{v^t B v}{v^t v} \geq \min_{x \in V} \frac{x^t B x}{x^t x} = \beta_k.$$

□

4 Gelfand's Formula

Lemma 4.1. *Let $A \in \mathbb{C}^{n \times n}$ have spectral radius $\rho(A)$. Then $\rho(A) < 1$ if and only if $A^k \to 0$. On the other hand, if $\rho(A) > 1$, then $\|A^k\| \to \infty$ for any choice of norm on $\mathbb{C}^{n \times n}$.*

Proof. Suppose $A^k \to 0$. We then have for any eigenvalue-eigenvector pair (λ, v),

$$0 = \lim_{k \to \infty} A^k v = \lim_{k \to \infty} \lambda^k v.$$

We must then have $|\lambda| < 1$. SInce this holds for any eigenvalue of A, we must have $\rho(A) < 1$. □

Theorem 4.2 (Gelfand's formula)**.** *If A is any any $n \times n$ matrix and $\| \cdot \|$ is any norm on $\mathbb{R}^{n \times n}$, then*

$$\rho(A) = \lim_{k \to \infty} \|A^k\|^{1/k},$$

where $\rho(A) = \max\{|\lambda| : \lambda$ is an eigenvalue of $A\}$ is the spectral radius of A.

Proof. The eigenvalues of A^k are simply the eigenvalues of A raised to the k-th power, so

$$\rho(A)^k = \rho(A^k) \leq \|A^k\| \implies \rho(A) \leq \|A^k\|^{1/k}.$$

□

5 The Perron-Frobenius Theorem

Definition 5.1. We say that a matrix is **elementwise nonnegative (positive)** if each of its entries is nonnegative (positive). We also write $A \geq_e B$ if $A - B$ is elementwise nonnegative.

Lemma 5.1. *A matrix $A \in \mathbb{R}^{m \times n}$ is elementwise nonnegative if $Ax \geq_e 0$ for all $x \geq_e 0$.*

Proof. If $A, x \geq_e 0$, then the entries of Ax are sums of nonnegative numbers, so $Ax \geq_e 0$. Conversely, if $Ax \geq_e 0$ for all $x \geq_e 0$, then $Ae_i \geq_e 0$ for all $1 \leq i \leq n$, where e_i is the vector in $\mathbb{R}^n$ with a 1 in the i-th slot and a zero everywhere else. Since Ae_i is the i-th column of A, we have that each column of A is elementwise nonnegative, so $A \geq_e 0$. □

Lemma 5.2. *Let $A >_e 0$ be an $n \times n$ matrix. If $u, v \in \mathbb{R}^n$ are unequal and $u \geq_e v$, then $Au >_e Av$. There is some $\epsilon > 0$ such that $Au >_e (1+\epsilon)Av$.*

Proof. The i-th entry of $A(u-v)$ is given by

$$[A(u-v)]_i = \sum_{j=1}^{n} A_{i,j}(u_i - v_i) \geq \min_{i,j} A_{i,j} \sum_{j=1}^{n} (u_i - v_i) > 0.$$

This holds for all i, so we have $Au >_e Av$. Since $A(u-v)$ is elementwise positive, we can perturb it by some small amount and keep it elementwise positive. There is then some $\epsilon > 0$ so that $A(u-v) - \epsilon Av >_e 0$, which proves the second part. $\square$