Some proofs on matrices

Ranveer*

1 Cauchy–Schwarz inequality

Theorem 1.1. *For any two vectors* $u, v \in \mathbb{R}^n$

$$|u^T v| \leq ||u|| ||v||.$$

Proof. Let $\{v_1, v_2, \ldots v_n\}$ be the set of orthogonal vector $v_i \in \mathbb{R}^n$, forming a basis of $\mathbb{R}^n$, where $v_1 = v$ (we can always form such a set). Let us write u as a linear combination

$$u = \alpha_1 v + \alpha_2 v_2 + \ldots + \alpha_n v_n,$$

$\alpha_i \in \mathbb{R}$. Then as $v_i^T v_i \geq 0$,

$$u^T u = \sum_{i=1}^{n} \alpha_i^2 v_i^T v_i \geq \alpha_1^2 v^T v.$$

Hence,

$$|\alpha_1| \leq \sqrt{\frac{u^T u}{v^T v}}.$$

Also, $v^T u = \alpha_1 v^T v$. Thus

$$|v^T u| = |\alpha_1| v^T v \leq \sqrt{\frac{u^T u}{v^T v}} v^T v = \sqrt{u^T u}\sqrt{v^T v} = ||u|| ||v||.$$

It completes the proof. □

2 The Spectral Theorem (Orthogonal diagonalization of symmteric matrices)

Theorem 2.1. *Let* M *be a symmetric matrix of order* n*. Then*

$$M = VDV^T, \textit{equivalently}, V^T M V = D,$$

where V *is a orthogonal matrix, (that is,* $V^T V = V V^T = I$*) with columns* $v_1, v_2, \ldots, v_n$ *which are the eigenvectors of* M *corresponding to the eigenvalues* $\lambda_1, \ldots, \lambda_n$*, and* $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$*.*

Proof. We prove the theorem using induction on n. The case $n = 1$ is true, since $M = 1M1^T$. Assume that the theorem is true when the order of matrix is $n - 1$.

Consider the case when the order of M is n. Using the eigenvector v_1, let us make an orthonormal basis of $\mathbb{R}^n$, suppose this be $S = \{v_1, x_2, \ldots, x_n\}$. Note that since M is symmetric, for any $i = 2, \ldots, i = n$

$$(Mx_i)^T v_1 = x_i^T M^T v_1 = x_i^T M v_1 = \lambda_1 x_i^T v_1 = 0. \quad (1)$$

*Email: ranveer@iiti.ac.in,

Now, let Q be an orthogonal matrix whose columns are $v_1, x_2, \ldots, x_n$. Then using 1

$$\begin{aligned} Q^T M Q &= \begin{bmatrix} - & v_1^T & - \\ - & x_2^T & - \\ \vdots & \vdots & \vdots \\ - & x_n^T & - \end{bmatrix} \begin{bmatrix} | & | & \ldots & | \\ \lambda_1 v_1 & M x_2 & \ldots & M x_n \\ | & | & \ldots & | \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \hat{M} \end{bmatrix}. \end{aligned}$$

Note that $\hat{M}$ is a symmetric matrix of order $n-1$. By induction hypothesis there exists an orthogonal matrix $\hat{Q}$ such that $\hat{Q}^T \hat{M} \hat{Q} = \hat{D}$ is a diagonal matrix. Considering the matrix

$$P = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q} \end{bmatrix},$$

we see that the product

$$\begin{aligned} P^T Q^T M Q P &= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q} \end{bmatrix}^T \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \hat{M} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q} \end{bmatrix} && (2) \\ &= \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \hat{D} \end{bmatrix} && (3) \end{aligned}$$

is a diagonal matrix. As $P^T Q^T Q P = I$, matrix QP is an orthogonal matrix. By (2) QP contains the eigenvectors of M and the corresponding eigenvalues are in the diagonal matrix $P^T Q^T M Q P$. It completes the proof. □

3 The largest and the smallest eigenvalue of a symmetric matrix (a nice optimization problem)

Theorem 3.1. *Let M be a symmetric matrix of order n with the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Then*

1. $\lambda_1 = \max_{||x||=1} x^T M x = \max_{x \neq \mathbf{0}} \frac{x^T M x}{||x||^2}$.
2. *If for some x we have $x^T M x = \lambda_1 ||x||^2$, then $M x = \lambda_1 x$*

Proof. Let $v_1, \ldots, v_n$ be unit eigenvectors corresponding to the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ of M.

1. Let us write x as a linear combination of $v_1, \ldots, v_n$,

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Let $V = [v_1, \ldots, v_n], \alpha = [\alpha_1, \ldots, \alpha_n]^T$, then $x = V\alpha$.
Then

$$||x||^2 = \alpha^T V^T V \alpha = \alpha^T \alpha = \sum_{i=1}^{n} \alpha_i^2,$$

and

$$x^T M x = \sum_{i=1}^{n} \lambda_i \alpha_i^2.$$

We see that

$$x^T M x = \sum_{i=1}^{n} \lambda_i \alpha_i^2 \leq \lambda_1 \sum_{i=1}^{n} \alpha_i^2 = \lambda_1 ||x||^2.$$

Since we know that

$$v_1^T M v_1 = \lambda_1 ||v_1||^2.$$

This implies

$$\lambda_1 = \max_{||x||=1} x^T Mx = \max_{x \neq \mathbf{0}} \frac{x^T Mx}{||x||^2}.$$

2. Consider $x^T Mx = \lambda_1 ||x||^2$. Then,

$$x^T Mx = \sum_{i=1}^{n} \lambda_i \alpha_i^2 = \lambda_1 \sum_{i=1}^{n} \alpha_i^2.$$

Suppose that $\lambda_1 = \cdots = \lambda_k > \lambda_{k+1} \geq \cdots \geq \lambda_n$. Then it implies that $\alpha_{k+1} = \cdots = \alpha_n = 0$, that is

$$x = \alpha_1 v_1 + \cdots + \alpha_k v_k.$$

Hence, $Mx = \lambda_1 x$.

□

Theorem 3.2. *Let M be a symmetric matrix of order n. The smallest eigenvalue λ_n of M is equal to*

$$\min_{||x||=1} x^T Mx = \min \frac{x^T Mx}{||x||^2}.$$

Proof. Let $v_1, \ldots, v_n$ be unit eigenvectors corresponding to the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ of M. Consider any vector x, and write it as a linear combination

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Then

$$x^T Mx = \sum_{i=1}^{n} \lambda_i \alpha_i^2 \geq \lambda_n \sum_{i=1}^{n} \alpha_i^2 = \lambda_n ||x||^2.$$

Since $Mv_n = \lambda_n v_n$,

$$\lambda_n = \min_{||x||=1} x^T Mx = \frac{x^T Mx}{||x||^2}.$$

□

4 Cauchy's Interlace Theorem

Theorem 4.1. *Let M be a symmetric matrix of order n, and B be its principal submatrix of order m, where, $m < n$. Suppose the eigenvalues of A are $\lambda_1 \geq \ldots \geq \lambda_n$ and the eigenvalues of B are $\beta_1 \geq \ldots \geq \beta_m$. Then*

$$\lambda_k \geq \beta_k \geq \lambda_{k+n-m}, \quad \textit{for } k = 1, \ldots, m.$$

Proof. For some suitable matrices X, Z we can write

$$M = \begin{bmatrix} B & X^T \\ X & Z \end{bmatrix}.$$

Let $x_1, \ldots, x_n$ be eigenvectors of A corresponding to the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_m$. And let $y_1, \ldots, y_n$ be eigenvectors of B corresponding to the eigenvalues $\beta_1 \geq \ldots \geq \beta_m$. Let us define the following two vector spaces.

1. $W = span(y_1, \ldots, y_k)$. We have
$$\beta_k = \min_{x \in W} \frac{x^T Bx}{x^T x} \tag{4}$$

2. $V = span(x_k, \ldots, x_n)$. We have
$$\lambda_k = \max_{x \in V} \frac{x^T Mx}{x^T x} \tag{5}$$

Construct a new vector space

$$\tilde{W} = \left\{ \begin{bmatrix} w \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^n \right\}, \ w \in W.$$

Since $\operatorname{rank} V = n - k + 1$, and $\operatorname{rank} \tilde{W} = k$, there exists $\tilde{w} = \begin{bmatrix} w \\ 0 \end{bmatrix} \in V \cap \tilde{W}$ for some $w \in W$. Then

$$\tilde{w}^T M \tilde{w} = \begin{bmatrix} w^T & 0 \end{bmatrix} \begin{bmatrix} B & X^T \\ X & Z \end{bmatrix} \begin{bmatrix} w \\ 0 \end{bmatrix} = w^T B w.$$

Using (4), (5)

$$\lambda_k \geq \frac{\tilde{w}^T M \tilde{w}}{\tilde{w}^T \tilde{w}} = \frac{w^T B w}{w^T w} \geq \beta_k.$$

The proof of the other inequality is similar. Now let us define the following two vector spaces.

1. $W = span(y_k, \ldots, y_m)$. Note that
$$\beta_k = \max_{x \in W} \frac{x^T B x}{x^T x} \tag{6}$$
2. $V = span(x_1, \ldots, x_{k+n-m})$. Note that
$$\lambda_{k+n-m} = \min_{x \in V} \frac{x^T M x}{x^T x} \tag{7}$$

Construct a new vector space

$$\tilde{W} = \left\{ \begin{bmatrix} w \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^n \right\}, \ w \in W.$$

Since $\operatorname{rank} V = k + n - m$, and $\operatorname{rank} \tilde{W} = m - k + 1$, there exists $\tilde{w} = \begin{bmatrix} w \\ 0 \end{bmatrix} \in V \cap \tilde{W}$ for some $w \in W$. Then as before we have $\tilde{w}^T M \tilde{w} = w^T B w$. Using (6), (7)

$$\lambda_{k+n-m} \leq \frac{\tilde{w}^T M \tilde{w}}{\tilde{w}^T \tilde{w}} = \frac{w^T B w}{w^T w} \leq \beta_k.$$

It completes the proof. □

For further references see Daniel A. Spielman, Embree.