Linear algebra

Prerequisites

Reduction of matrices

Theorem 1.2-1

(1) *Given a square matrix* A, *there exists a unitary matrix* U *such that the matrix* $U^{-1}AU$ *is triangular.*

(2) *Given a normal matrix* A, *there exists a unitary matrix* U *such that the matrix* $U^{-1}AU$ *is diagonal.*

(3) *Given a symmetric matrix* A, *there exists an orthogonal matrix* O *such that the matrix* $O^{-1}AO$ *is diagonal.*

Theorem 1.2-2

If A *is a real, square matrix, there exist two orthogonal matrices* U *and* V *such that*

$$U^T AV = \operatorname{diag}(\mu_i),$$

and, if A *is a complex, square matrix, there exist two unitary matrices* U *and* V *such that*

$$U^* AV = \operatorname{diag}(\mu_i).$$

In either case, the numbers $\mu_i \geqslant 0$ *are the singular values of the matrix* A.

Rayleigh quotient

Theorem 1.3-1

Let A *be a Hermitian matrix of order n, with eigenvalues*

$$\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n,$$

the associated eigenvectors $p_1, p_2, \ldots, p_n$ *satisfying*

$$p_i^* p_j = \delta_{ij}.$$

For $k = 1, \ldots, n$, *let* V_k *denote the subspace of* V *spanned by the vectors* p_i, $1 \leqslant i \leqslant k$, *and let* $\mathscr{V}_k$ *denote the set of subspaces of* V *of dimension* k. *Furthermore, set*

$$V_0 = \{0\}, \mathscr{V}_0 = \{V_0\}.$$

The eigenvalues then have the following characterisations, for $k = 1, 2, \ldots, n$.

(1) $$\lambda_k = R_A(p_k),$$

(2) $$\lambda_k = \max_{v \in V_k} R_A(v),$$

(3) $$\lambda_k = \min_{v \perp V_{k-1}} R_A(v),$$

(4) $$\lambda_k = \min_{W \in \mathscr{V}_k} \max_{v \in W} R_A(v),$$

(5) $$\lambda_k = \max_{W \in \mathscr{V}_{k-1}} \min_{v \perp W} R_A(v).$$

Furthermore,

(6) $$\{R_A(v) : v \in V\} = [\lambda_1, \lambda_n] \subset \mathbb{R}.$$

Norms upon matrice spaces

Theorem 1.4-2

Let $A = (a_{ij})$ be a square matrix. Then

$$\|A\|_1 \stackrel{\text{def}}{=} \sup \frac{\|Av\|_1}{\|v\|_1} = \max_j \sum_i |a_{ij}|,$$

$$\|A\|_2 \stackrel{\text{def}}{=} \sup \frac{\|Av\|_2}{\|v\|_2} = \sqrt{\varrho(A^*A)} = \sqrt{\varrho(AA^*)} = \|A^*\|_2,$$

$$\|A\|_\infty \stackrel{\text{def}}{=} \sup \frac{\|Av\|_\infty}{\|v\|_\infty} = \max_i \sum_j |a_{ij}|.$$

The norm $\|\cdot\|_2$ is invariant under unitary transformations:

$$UU^* = I \Rightarrow \|A\|_2 = \|AU\|_2 = \|UA\|_2 = \|U^*AU\|_2.$$

Furthermore, if the matrix A is normal,

$$AA^* = AA^* \Rightarrow \|A\|_2 = \varrho(A).$$

Theorem 1.4-3

(1) *Let* A *be any square matrix and* $\|\cdot\|$ *any matrix norm, subordinate or otherwise. Then*

$$\varrho(\mathrm{A}) \leqslant \|\mathrm{A}\|.$$

(2) *Given a matrix* A *and any number* $\varepsilon > 0$, *there exists at least one subordinate matrix norm such that*

$$\|\mathrm{A}\| \leqslant \varrho(\mathrm{A}) + \varepsilon.$$

Theorem 1.4-4

The function $\|\cdot\|_{\mathrm{E}}\colon \mathscr{A}_n \to \mathbb{R}$ *defined by*

$$\|\mathrm{A}\|_{\mathrm{E}} = \left\{\sum_{i,j} |a_{ij}|^2\right\}^{1/2} = \{\operatorname{tr}(\mathrm{A}^*\mathrm{A})\}^{1/2}$$

for every matrix $\mathrm{A} = (a_{ij})$ *of order* n *is a matrix norm which is not subordinate* (*for* $n \geqslant 2$), *is invariant under unitary transformations,*

$$\mathrm{UU}^* = \mathrm{I} \Rightarrow \|\mathrm{A}\|_{\mathrm{E}} = \|\mathrm{AU}\|_{\mathrm{E}} = \|\mathrm{UA}\|_{\mathrm{E}} = \|\mathrm{U}^*\mathrm{AU}\|_{\mathrm{E}},$$

and satisfies

$$\|\mathrm{A}\|_2 \leqslant \|\mathrm{A}\|_{\mathrm{E}} \leqslant \sqrt{n}\,\|\mathrm{A}\|_2 \quad \text{for every } A \in \mathscr{A}_n.$$

Theorem 1.4-5

(1) *Let* $\|\cdot\|$ *be a subordinate matrix norm and* B *a matrix satisfying*

$$\|\mathbf{B}\| < 1.$$

Then the matrix $\mathrm{I} + \mathrm{B}$ *is invertible and*

$$\|(\mathrm{I} + \mathrm{B})^{-1}\| \leqslant \frac{1}{1 - \|\mathbf{B}\|}.$$

(2) *If a matrix of the form* $\mathrm{I} + \mathrm{B}$ *is singular, then necessarily*

$$\|\mathbf{B}\| \geqslant 1$$

for every matrix norm, subordinate or not.

Convergence of a suite of matrice

Theorem 1.5-1

Let B *be a square matrix. The following conditions are equivalent:*

(1) $$\lim_{k \to \infty} \mathrm{B}^k = 0,$$

(2) $$\lim_{k \to \infty} \mathrm{B}^k v = 0 \textit{ for every vector } v,$$

(3) $$\varrho(\mathrm{B}) < 1,$$

(4) $$\| \mathrm{B} \| < 1 \textit{ for at least one subordinate matrix norm } \| \cdot \|.$$

Theorem 1.5-2

Let B *be a square matrix and let* $\| \cdot \|$ *be any matrix norm. Then*

$$\lim_{k \to \infty} \| \mathrm{B}^k \|^{1/k} = \varrho(\mathrm{B}).$$

Conditioning

Theorem 2.2-1
Let A *be an invertible matrix and let u and $u + \delta u$ be the solutions of the linear systems*

$$\mathrm{A}u = b,$$
$$\mathrm{A}(u + \delta u) = b + \delta b.$$

Suppose that $b \neq 0$. Then the inequality

$$\frac{\|\delta u\|}{\|u\|} \leqslant \mathrm{cond}(\mathrm{A}) \frac{\|\delta b\|}{\|b\|}$$

holds and is the best possible: that is, for a given matrix A, *it is possible to find vectors $b \neq 0$ and $\delta b \neq 0$ for which equality holds.*

Theorem 2.2-3

(1) *For every matrix* A,

$$\operatorname{cond}(A) \geqslant 1,$$

$$\operatorname{cond}(A) = \operatorname{cond}(A^{-1}),$$

$$\operatorname{cond}(\alpha A) = \operatorname{cond}(A) \quad \textit{for every scalar } \alpha \neq 0.$$

(2) *For every matrix* A,

$$\operatorname{cond}_2(A) = \frac{\mu_n(A)}{\mu_1(A)},$$

where $\mu_1(A) > 0$ *and* $\mu_n(A) > 0$ *denote respectively the smallest and the largest of the singular values of the matrix* A.

(3) *If the matrix A is normal*

$$\operatorname{cond}_2(A) = \frac{\max_i |\lambda_i(A)|}{\min_i |\lambda_i(A)|},$$

where the numbers $\lambda_i(A)$ *are the eigenvalues of the matrix* A.

(4) *The condition number* $\operatorname{cond}_2(A)$ *of a unitary or orthogonal matrix* A *is equal to* 1.

(5) *The condition number* $\operatorname{cond}_2(A)$ *is invariant under unitary transformations:*

$$UU^* = I \Rightarrow \operatorname{cond}_2(A) = \operatorname{cond}_2(AU) = \operatorname{cond}_2(UA) = \operatorname{cond}_2(U^*AU).$$

Theorem 2.2-2

Let A *be an invertible matrix and let* u *and* $u + \Delta u$ *be the solutions of the linear systems*

$$\mathrm{A}u = b,$$

$$(\mathrm{A} + \Delta\mathrm{A})(u + \Delta u) = b.$$

Suppose that $b \neq 0$. *Then the inequality*

$$\frac{\|\Delta u\|}{\|u + \Delta u\|} \leqslant \operatorname{cond}(\mathrm{A}) \frac{\|\Delta\mathrm{A}\|}{\|\mathrm{A}\|}$$

holds and is the best possible: that is, for a given matrix A, *it is possible to find a vector* $b \neq 0$ *and a matrix* $\Delta\mathrm{A} \neq 0$ *for which equality holds.*

Furthermore, we have the inequality

$$\frac{\|\Delta u\|}{\|u\|} \leqslant \operatorname{cond}(\mathrm{A}) \frac{\|\Delta\mathrm{A}\|}{\|\mathrm{A}\|}\{1 + \mathrm{O}(\|\Delta\mathrm{A}\|)\}.$$