CS730
Appendix A
Vector and Matrix Norms

1. Vector norm (Def.)

A function $\nu : R^n \to R$ is a vector norm on R^n iff:

(a) $\nu(x) \geqq 0 \qquad \forall x \in R^n$ and $\nu(x) = 0$ iff $x = 0$

(b) $\nu(\alpha x) = |\alpha| \nu(x) \qquad \forall x \in R^n,\ \forall \alpha \in R$

(c) $\nu(x + y) \leqq \nu(x) + \nu(y) \qquad \forall x,\ y \in R^n$ (Triangle Inequality)

Usually $\nu(x)$ is denoted by $\|x\|$.

Examples:

$$\|x\|_2 = \Big(\sum_{i=1}^{n} x_i^2\Big)^{\frac{1}{2}} \qquad (\ell_2 \text{ or Euclidean})$$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \qquad (\ell_1)$$

$$\|x\|_\infty = \max_1 \leqq i \leqq n\, |x_1| \qquad (\ell_\infty \text{ or max norm})$$

$$\|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}}\ 1 \leqq p < \infty \qquad (\ell p)$$

2. Norm equivalence theorem

For any two norms: μ and ν on R^n, there exist positive numbers $\alpha_{\nu\mu}$ and $\beta_{\nu\mu}$ such that

$$\alpha_{\nu\mu} \cdot \mu(x) \leqq \nu(x) \leqq \nu\beta_\nu\mu \cdot \mu(x) \qquad \forall x \in R^n,$$

with equalities holding for some $x \in R^n$. For example:

$$\|x\|_\infty \leqq \|x\|_2 \leqq \|x\|_1 \leqq n^{\frac{1}{2}} \|x\|_2 \leqq n\|x\|_\infty$$

3. Dual norm (Def.)

The function $\|\ \ \|' : R^n \to R$ defined by

$$\|x\|' = \max_{y \in R^n,\, \|y\|=1} xy$$

for any norm $\|\ \ \|$ on R^n, is also a norm on R^n and is called the dual norm of $\|\ \ \|$.

4. Generalized Cauchy inequality

$$|xy| \leqq \|x\|' \cdot \|y\| \qquad \forall x,\ y \in R^n$$

Proof. $\|x\|' = \max\limits_{\|y\|=1} |xy| = \max\limits_{y \neq 0} \left| x \dfrac{y}{\|y\|} \right|$.

Hence $\|x\|' \geqq \left| x \dfrac{y}{\|y\|} \right|$ for $\forall y \neq 0$ and $|xy| \leqq \|x\|' \cdot \|y\|$ for all $x,\ y$ in R^n. □

5. Dual ℓ_p norms theorem

The norms $\|\cdot\|_p$ and $\|\cdot\|_q$ are dual norms for $1 \leqq p \leqq \infty$ and $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

6. Hölder inequality

$$|xy| \leqq \|x\|_p \|y\|_q, \text{ for } \frac{1}{p} + \frac{1}{q} = 1, \; 1 \leqq p \leqq \infty$$

7. Cauchy inequality

$$|xy| \leqq \|x\|_2 \|y\|_2$$

8. Matrix norm (Operator norm) (Def.)

A function $\|\cdot\| : R^{m\times n} \to R$ is a matrix norm iff

$$\|A\| = \max_{\substack{\|x\|=1 \\ x\in R^n}} \|Ax\| = \max_{\substack{x\neq 0 \\ x\in R^n}} \frac{\|Ax\|}{\|x\|}$$

9. $\|\mathbf{Ax}\| \leqq \|\mathbf{A}\| \;\; \|\mathbf{x}\|$

Proof. From the definition of a matrix norm $\|A\| \geqq \dfrac{\|Ax\|}{\|x\|}$ for all $x \in R^n, \; x \neq 0$ and hence $\|A\| \;\; \|x\| \geqq \|Ax\|$ for all $x \in R^n$. $\square$

10. Properties of matrix norms

(a) $\|A\| \geqq 0$; $\|A\| = 0$ iff $A = 0$

(b) $\|\alpha A\| = |\alpha| \cdot \|A\|, \; \alpha \in R$

(c) $\|A + B\| \leqq \|A\| + \|B\|$

(These 3 properties are sometimes used as an alternate definition of a matrix norm that is not an operator norm).

(d) $\|AB\| \leqq \|A\| \;\; \|B\|$

Consistency condition not necessarily satisfied by norms other than operator norms.

11. Examples of matrix norms ($\mathbf{A} \in \mathbf{R}^{\mathbf{n}\times\mathbf{n}}$)

$\|A\|_2 = \max\limits_{\|x\|_2=1} \|Ax\|_2$

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max \text{ "row sum" } = \max_{1 \leqq i \leqq n} \sum_{i=1}^{n} |a_{ij}|$$

$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max \text{ "column sum" } = \max_{1 \leqq j \leqq n} \sum_{j=1}^{n} |a_{ij}|$$

12. Spectral radius $\rho(\mathbf{A})$: ($\mathbf{A} \in \mathbf{R}^{\mathbf{n}\times\mathbf{n}}$)

$$\rho(A) = \max\left\{|\lambda| \,\Big|\, Ax = \lambda x, \; x \neq 0\right\}$$

13. Theorem

Let $A \in R^{n \times n}$. Then:

$$\rho(A) = \inf_{\nu} \left\{ \|A\|_\nu \Big| \| \cdot \|_\nu = \text{an operator norm on } R^{n \times n} \right\}$$

That is

(a) $\rho(A) \leqq \|A\|_\nu$ for any operator norm on $R^{n \times n}$

(b) For each $\varepsilon > 0$ there exists an operator norm $\| \quad \|_\varepsilon$ such that

$$\rho(A) + \varepsilon > \|A\|_\varepsilon$$

14. Theorem

Let $\lambda_1 \leqq, \ldots, \leqq \lambda_n$ be the real eigenvalues of the symmetric matrix $A \in R^{n \times n}$. Then

$$\lambda_1 xx \leqq xAx \leqq \lambda_n xx \qquad \forall x \in R^n$$

Proof.
Let P be the orthogonal matrix of the eigenvectors of A. Then $P^T AP = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and $P^T P = I$. Hence, with $y = P^T x$ and $x = Py$, $xAx = yP^T APy = \sum_{i=1}^{n} \lambda_i y_i^2$.

So

$$\lambda_1 xx = \lambda_1 yy \leqq xAx = \sum_{i=1}^{n} \lambda_i y_i^2 \leqq \lambda_n yy = \lambda_n xx.$$

□

15. Theorem

Let $A \in R^{n \times n}$, then

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2 = \Big(\rho(A^T A)\Big)^{\frac{1}{2}} \qquad (*)$$

Proof. Define $\mu = \Big(\rho(A^T A)\Big)^{\frac{1}{2}}$. Then by Theorem 14

$$\|Ax\|_2^2 = xA^T Ax \leqq \mu^2 x^T x$$

so that $\|A\|_2 \leqq \mu$. On the other hand, if u is an eigenvector of $A^T A$ corresponding to μ^2 then

$$uA^T Au = \mu^2 uu$$

which shows that equality holds in $(*)$. □

16. Theorem

Let $A \in R^{n \times n}$, then

$$\lim_{k \to \infty} A^k = 0 \Longleftrightarrow \rho(A) < 1$$

17. Neumann Lemma

$$A \in R^{n \times n},\ \rho(A) < 1 \Longrightarrow (I - A)^{-1} = \sum_{i=0}^{\infty} A^i$$

18. Banach Perturbation Lemma

$$\left.\begin{array}{c} A \in R^{n \times n}, E \in R^{n \times n} (\text{E=perturbation}) \\ A^{-1} \text{ exists} \\ \|A^{-1}\|\ \|E\| < 1 \end{array}\right\rangle \Longrightarrow \left\langle \begin{array}{c} (A+E)^{-1} \text{exists and:} \\ \dfrac{1}{\|A+E\|} \leqq \|(A+E)^{-1}\| \leqq \dfrac{\|A^{-1}\|}{1 - \|A^{-1}\|\ \|E\|} \end{array}\right.$$

Proof. Set $-B = A^{-1}E$. Then $\|B\| < 1$ and by the Neumann Lemma, $(I - B)^{-1} = \sum_{i=0}^{\infty} B^i$. Hence

$$\|(I - B)^{-1}\| \leqq \sum_{i=0}^{\infty} \|B\|^i = \frac{1}{1 - \|B\|} \leqq \frac{1}{1 - \|A^{-1}\|\ \|E\|}.$$

But

$$A + E = A(I + A^{-1}E) = A(I - B),$$

so that $A + E$, the product of nonsingular matrices, is nonsingular and

$$\|(A+E)^{-1}\| = \|(I-B)^{-1}A^{-1}\| \leqq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\ \|E\|}.$$

The lower bound of the lemma easily follows from:

$$1 = \|I\| = \|(A+E)(A+E))^{-1}\| \leqq \|A+E\|\ \|(A+E)^{-1}\|.$$

$\square$

19. Condition number (Def.)

Let $A \in R^{n \times n}$ then

$$K(A) = \begin{cases} \|A\|\ \|A^{-1}\| & \text{if } A \text{ is nonsingular} \\ \infty & \text{if } A \text{ is singular} \end{cases}$$

20. Error estimate

Let $A \in R^{n \times n}$ be nonsingular and let

$$Ax = b \text{ and } Ay = b + e,$$

where e is the error caused by replacing x by y. Then

$$x - y = -A^{-1}e \text{ and } \|A\| \cdot \|x\| \geqq \|b\|.$$

So

$$\frac{\|x-y\|}{\|x\|} \leqq \frac{\|A^{-1}e\|}{\|A\|^{-1}\|b\|} \leqq \|A\|\ \|A^{-1}\| \frac{\|e\|}{\|b\|} = K(A)\frac{\|e\|}{\|b\|}$$

21. Corollary to Banach Perturbation Lemma

$$\left.\begin{array}{c} A \in R^{n\times n},\ E \in R^{n\times n} \\ A^{-1}\text{exists} \\ \|A^{-1}\|\,\|E\| < 1 \end{array}\right\rangle \Longrightarrow \left\langle \begin{array}{rl} \|(A+E)^{-1} - A^{-1}\| & \leqq \dfrac{\|A^{-1}\|^2\|E\|}{1-\|A^{-1}\|\,\|E\|} \\ & \leqq \dfrac{\|A^{-1}\|}{1-\|A^{-1}\|\,\|E\|} \end{array}\right.$$

Proof. From the identity

$$(A+E)^{-1} - A^{-1} = -A^{-1}E(A+E)^{-1},$$

we get, using the Banach perturbation Lemma 18, that:

$$\|(A+E)^{-1} - A^{-1}\| \leqq \frac{\|A^{-1}\|^2\|E\|}{1-\|A^{-1}\|\,\|E\|} \leqq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\,\|E\|}.$$

□

22. Error estimate for perturbed equations

$$\left.\begin{array}{l} A\in R^{n\times n}, E\in R^{n\times n}(\text{E=perturbation}) \\ A^{-1}\text{ exists} \\ \|A^{-1}\|\,\|E\| < 1 \\ Ax = b,\ b \in R^n \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \text{For each perturbation } e \in R^n \text{ of } b, \\ (A+E)\tilde{x} = b + e \text{ has a unique solution} \\ \tilde{x} = (A+E)^{-1}(b+e) \text{ satisfying :} \\ \|\tilde{x} - x\| \leqq \dfrac{\|A^{-1}\|}{1-\|A^{-1}\|\,\|E\|}[\|e\| + \|A^{-1}\|\,\|E\|\,\|b\|] \\ \qquad\qquad \leqq \dfrac{\|A^{-1}\|}{1-\|A^{-1}\|\,\|E\|}[\|e\| + \|b\|] \end{array}\right.$$

Proof. The above estimates follow by applying the Banach perturbation Lemma 18 and its Corollary 21 to

$$\tilde{x} - x = (A+E)^{-1}e + \Big((A+E)^{-1} - A^{-1}\Big)b$$

.

□

References

1. J. S. Ortega: "Numerical analysis a second course", Academic Press, New York, 1972.
2. G. W. Stewart: "Introduction to matrix computations", Academic Press, New York, 1973.