11.2   Norms and Condition Numbers

How do we measure the size of a matrix? For a vector, the length is $\|x\|$. For a matrix, *the norm is* $\|A\|$. This word "*norm*" is sometimes used for vectors, instead of length. It is always used for matrices, and there are many ways to measure $\|A\|$. We look at the requirements on all "matrix norms" and then choose one.

Frobenius squared all the $|a_{ij}|^2$ and added; his norm $\|A\|_{\mathrm{F}}$ is the square root. This treats A like a long vector with n^2 components: sometimes useful, but not the choice here.

I prefer to start with a vector norm. The triangle inequality says that $\|x + y\|$ is not greater than $\|x\| + \|y\|$. The length of $2x$ or $-2x$ is doubled to $2\|x\|$. The same rules will apply to matrix norms:

$$\|A + B\| \le \|A\| + \|B\| \qquad \text{and} \qquad \|cA\| = |c|\,\|A\|. \tag{1}$$

The second requirements for a matrix norm are new, because matrices multiply. The norm $\|A\|$ controls the growth from x to Ax, and from B to AB:

Growth factor $\|A\|$ $\qquad \|Ax\| \le \|A\|\,\|x\| \qquad \text{and} \qquad \|AB\| \le \|A\|\,\|B\|. \tag{2}$

This leads to a natural way to define $\|A\|$, the norm of a matrix:

$$\textit{\textbf{The norm of A is the largest ratio }} \|Ax\|/\|x\|: \qquad \|A\| = \max_{x \ne 0} \frac{\|Ax\|}{\|x\|}. \tag{3}$$

$\|Ax\|/\|x\|$ is never larger than $\|A\|$ (its maximum). This says that $\|Ax\| \le \|A\|\,\|x\|$.

Example 1   If A is the identity matrix I, the ratios are $\|x\|/\|x\|$. Therefore $\|I\| = 1$. If A is an orthogonal matrix Q, lengths are again preserved: $\|Qx\| = \|x\|$. The ratios still give $\|Q\| = 1$. An orthogonal Q is good to compute with: errors don't grow.

Example 2   The norm of a diagonal matrix is its largest entry (using absolute values):

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{has norm} \quad \|A\| = 3. \quad \text{The eigenvector} \quad x = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{has} \quad Ax = 3x.$$

The eigenvalue is 3. For this A (but not all A), the largest eigenvalue equals the norm.

*$\qquad$ **For a positive definite symmetric matrix the norm is** $\|A\| = \lambda_{\max}(A)$.*

Choose x to be the eigenvector with maximum eigenvalue. Then $\|Ax\|/\|x\|$ equals $\lambda_{\max}$. The point is that no other x can make the ratio larger. The matrix is $A = Q\Lambda Q^{\mathrm{T}}$, and the

orthogonal matrices Q and Q^{T} leave lengths unchanged. So the ratio to maximize is really $\|\Lambda x\|/\|x\|$. The norm is the largest eigenvalue in the diagonal Λ.

Symmetric matrices Suppose A is symmetric but not positive definite. $A = Q\Lambda Q^{\mathrm{T}}$ is still true. Then the norm is the largest of $|\lambda_1|$, $|\lambda_2|$, ..., $|\lambda_n|$. We take absolute values, because the norm is only concerned with length. For an eigenvector $\|Ax\| = \|\lambda x\| = |\lambda|$ times $\|x\|$. The x that gives the maximum ratio is the eigenvector for the maximum $|\lambda|$.

Unsymmetric matrices If A is not symmetric, its eigenvalues may not measure its true size. *The norm can be larger than any eigenvalue.* A very unsymmetric example has $\lambda_1 = \lambda_2 = 0$ but its norm is not zero:

$$\|A\| > \lambda_{\mathbf{max}} \quad A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \qquad \text{has norm} \qquad \|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = 2.$$

The vector $x = (0, 1)$ gives $Ax = (2, 0)$. The ratio of lengths is $2/1$. This is the maximum ratio $\|A\|$, even though x is not an eigenvector.

It is the ***symmetric matrix*** $A^{\mathrm{T}}A$, not the unsymmetric A, that has eigenvector $x = (0, 1)$. The norm is really decided by ***the largest eigenvalue of*** $A^{\mathrm{T}}A$:

The norm of A (symmetric or not) ***is the square root of*** $\lambda_{\mathbf{max}}(A^{\mathrm{T}}A)$***:***

$$\|A\|^2 \;\; = \max_{x \neq 0} \frac{\|Ax\|^2}{\|x\|^2} = \max_{x \neq 0} \frac{x^{\mathrm{T}}A^{\mathrm{T}}Ax}{x^{\mathrm{T}}x} = \;\; \lambda_{\max}(A^{\mathrm{T}}A) \;. \tag{4}$$

The unsymmetric example with $\lambda_{\max}(A) = 0$ has $\lambda_{\max}(A^{\mathrm{T}}A) = 4$:

$$A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \text{ leads to } A^{\mathrm{T}}A = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \text{ with } \lambda_{\max} = 4. \text{ So the norm is } \|A\| = \sqrt{4}.$$

For any A Choose x to be the eigenvector of $A^{\mathrm{T}}A$ with largest eigenvalue $\lambda_{\max}$. The ratio in equation (4) is $x^{\mathrm{T}}A^{\mathrm{T}}Ax = x^{\mathrm{T}}(\lambda_{\max})x$ divided by $x^{\mathrm{T}}x$. This is $\lambda_{\max}$.

No x can give a larger ratio. The symmetric matrix $A^{\mathrm{T}}A$ has eigenvalues $\lambda_1, \ldots, \lambda_n$ and orthonormal eigenvectors $q_1, q_2, \ldots, q_n$. Every x is a combination of those vectors. Try this combination in the ratio and remember that $q_i^{\mathrm{T}}q_j = 0$:

$$\frac{x^{\mathrm{T}}A^{\mathrm{T}}Ax}{x^{\mathrm{T}}x} = \frac{(c_1 q_1 + \cdots + c_n q_n)^{\mathrm{T}}(c_1\lambda_1 q_1 + \cdots + c_n\lambda_n q_n)}{(c_1 q_1 + \cdots + c_n q_n)^{\mathrm{T}}(c_1 q_1 + \cdots + c_n q_n)} = \frac{c_1^2\lambda_1 + \cdots + c_n^2\lambda_n}{c_1^2 + \cdots + c_n^2}.$$

The maximum ratio $\lambda_{\max}$ is when all c's are zero, except the one that multiplies $\lambda_{\max}$.

Note 1 The ratio in equation (4) is the ***Rayleigh quotient*** for the symmetric matrix $A^{\mathrm{T}}A$. Its maximum is the largest eigenvalue $\lambda_{\max}(A^{\mathrm{T}}A)$. The minimum ratio is $\lambda_{\min}(A^{\mathrm{T}}A)$.

If you substitute any vector x into the Rayleigh quotient $x^T A^T Ax/x^T x$, you are guaranteed to get a number between $\lambda_{\min}(A^T A)$ and $\lambda_{\max}(A^T A)$.

Note 2 The norm $\|A\|$ equals the ***largest singular value*** $\sigma_{\mathbf{max}}$ of A. The singular values $\sigma_1, \ldots, \sigma_r$ are the square roots of the positive eigenvalues of $A^T A$. So certainly $\sigma_{\max} = (\lambda_{\max})^{1/2}$. Since U and V are orthogonal in $A = U\Sigma V^T$, the norm is $\|\mathbf{A}\| = \sigma_{\mathbf{max}}$.

The Condition Number of A

Section 9.1 showed that roundoff error can be serious. Some systems are sensitive, others are not so sensitive. The sensitivity to error is measured by the ***condition number***. This is the first chapter in the book which intentionally introduces errors. We want to estimate how much they change x.

The original equation is $Ax = b$. Suppose the right side is changed to $b + \Delta b$ because of roundoff or measurement error. The solution is then changed to $x + \Delta x$. Our goal is to estimate the change Δx in the solution from the change Δb in the equation. Subtraction gives the *error equation* $A(\Delta x) = \Delta b$:

$$\text{Subtract } Ax = b \text{ from } A(x + \Delta x) = b + \Delta b \quad \text{to find} \quad \boxed{A(\Delta x) = \Delta b.} \quad (5)$$

The error is $\Delta x = A^{-1}\Delta b$. It is large when A^{-1} is large (then A is nearly singular). The error Δx is especially large when Δb points in the worst direction—which is amplified most by A^{-1}. ***The worst error has*** $\|\Delta x\| = \|A^{-1}\| \, \|\Delta b\|$.

This error bound $\|A^{-1}\|$ has one serious drawback. If we multiply A by 1000, then A^{-1} is divided by 1000. The matrix looks a thousand times better. But a simple rescaling cannot change the reality of the problem. It is true that Δx will be divided by 1000, but so will the exact solution $x = A^{-1}b$. The ***relative error*** $\|\Delta x\|/\|x\|$ will stay the same. It is this relative change in x that should be compared to the relative change in b.

Comparing relative errors will now lead to the "condition number" $c = \|A\| \, \|A^{-1}\|$. Multiplying A by 1000 does not change this number, because A^{-1} is divided by 1000 and the condition number c stays the same. It measures the sensitivity of $Ax = b$.

The solution error is less than $c = \|A\| \, \|A^{-1}\|$ ***times the problem error:***

Condition number c
$$\frac{\|\Delta x\|}{\|x\|} \le c \frac{\|\Delta b\|}{\|b\|}. \quad (6)$$

If the problem error is ΔA (error in A instead of b), still c controls Δx:

Error ΔA in A
$$\frac{\|\Delta x\|}{\|x + \Delta x\|} \le c \frac{\|\Delta A\|}{\|A\|}. \quad (7)$$

Proof The original equation is $b = Ax$. The error equation (5) is $\Delta x = A^{-1}\Delta b$. Apply the key property $\|Ax\| \leq \|A\|\|x\|$ of matrix norms:

$$\|b\| \leq \|A\|\,\|x\| \qquad \text{and} \qquad \|\Delta x\| \leq \|A^{-1}\|\,\|\Delta b\|.$$

Multiply the left sides to get $\|b\|\,\|\Delta x\|$, and multiply the right sides to get $c\|x\|\,\|\Delta b\|$. Divide both sides by $\|b\|\,\|x\|$. The left side is now the relative error $\|\Delta x\|/\|x\|$. The right side is now the upper bound in equation (6).

The same condition number $c = \|A\|\,\|A^{-1}\|$ appears when the error is in the matrix. We have ΔA instead of Δb in the error equation:

Subtract $Ax = b$ from $(A + \Delta A)(x + \Delta x) = b$ to find $A(\Delta x) = -(\Delta A)(x + \Delta x)$.

Multiply the last equation by A^{-1} and take norms to reach equation (7):

$$\|\Delta x\| \leq \|A^{-1}\|\,\|\Delta A\|\,\|x + \Delta x\| \quad \text{or} \quad \frac{\|\Delta x\|}{\|x + \Delta x\|} \leq \|A\|\,\|A^{-1}\|\frac{\|\Delta A\|}{\|A\|}.$$

Conclusion Errors enter in two ways. They begin with an error ΔA or Δb—a wrong matrix or a wrong b. This problem error is amplified (a lot or a little) into the solution error Δx. That error is bounded, relative to x itself, by the condition number c.

The error Δb depends on computer roundoff and on the original measurements of b. The error ΔA also depends on the elimination steps. Small pivots tend to produce large errors in L and U. Then $L + \Delta L$ times $U + \Delta U$ equals $A + \Delta A$. When ΔA or the condition number is very large, the error Δx can be unacceptable.

Example 3 When A is symmetric, $c = \|A\|\,\|A^{-1}\|$ comes from the eigenvalues:

$$A = \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix} \text{ has norm } 6. \qquad A^{-1} = \begin{bmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \text{ has norm } \frac{1}{2}.$$

This A is symmetric positive definite. Its norm is $\lambda_{\max} = 6$. The norm of A^{-1} is $1/\lambda_{\min} = \frac{1}{2}$. Multiplying norms gives the *condition number* $\|A\|\,\|A^{-1}\| = \lambda_{\max}/\lambda_{\min}$:

Condition number for positive definite A $\qquad c = \dfrac{\lambda_{\max}}{\lambda_{\min}} = \dfrac{6}{2} = 3.$

Example 4 Keep the same A, with eigenvalues 6 and 2. To make x small, choose b along the first eigenvector $(1, 0)$. To make Δx large, choose Δb along the second eigenvector $(0, 1)$. Then $x = \frac{1}{6}b$ and $\Delta x = \frac{1}{2}\Delta b$. The ratio $\|\Delta x\|/\|x\|$ is exactly $c = 3$ times the ratio $\|\Delta b\|/\|b\|$.

This shows that the worst error allowed by the condition number $\|A\|\,\|A^{-1}\|$ can actually happen. Here is a useful rule of thumb, experimentally verified for Gaussian elimination: *The computer can lose* $\log c$ *decimal places to roundoff error.*

Problem Set 11.2

1 Find the norms $\|A\| = \lambda_{\max}$ and condition numbers $c = \lambda_{\max}/\lambda_{\min}$ of these positive definite matrices:

$$\begin{bmatrix} .5 & 0 \\ 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}.$$

2 Find the norms and condition numbers from the square roots of $\lambda_{\max}(A^{\mathrm{T}}A)$ and $\lambda_{\min}(A^{\mathrm{T}}A)$. Without positive definiteness in A, we go to $A^{\mathrm{T}}A$!

$$\begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

3 Explain these two inequalities from the definitions (3) of $\|A\|$ and $\|B\|$:

$$\|ABx\| \le \|A\| \, \|Bx\| \le \|A\| \, \|B\| \, \|x\|.$$

From the ratio of $\|ABx\|$ to $\|x\|$, deduce that $\|AB\| \le \|A\| \, \|B\|$. This is the key to using matrix norms. The norm of A^n is never larger than $\|A\|^n$.

4 Use $\|AA^{-1}\| \le \|A\| \, \|A^{-1}\|$ to prove that the condition number is at least 1.

5 Why is I the only symmetric positive definite matrix that has $\lambda_{\max} = \lambda_{\min} = 1$? Then the only other matrices with $\|A\| = 1$ and $\|A^{-1}\| = 1$ must have $A^{\mathrm{T}}A = I$. Those are _____ matrices: perfectly conditioned.

6 Orthogonal matrices have norm $\|Q\| = 1$. If $A = QR$ show that $\|A\| \le \|R\|$ and also $\|R\| \le \|A\|$. Then $\|A\| = \|Q\| \, \|R\|$. Find an example of $A = LU$ with $\|A\| < \|L\| \|U\|$.

7 (a) Which famous inequality gives $\|(A + B)x\| \le \|Ax\| + \|Bx\|$ for every x?

  (b) Why does the definition (3) of matrix norms lead to $\|A + B\| \le \|A\| + \|B\|$?

8 Show that if λ is any eigenvalue of A, then $|\lambda| \le \|A\|$. Start from $Ax = \lambda x$.

9 The "*spectral radius*" $\rho(A) = |\lambda_{\max}|$ is the largest absolute value of the eigenvalues. Show with 2 by 2 examples that $\rho(A + B) \le \rho(A) + \rho(B)$ and $\rho(AB) \le \rho(A)\rho(B)$ can both be *false*. The spectral radius is not acceptable as a norm.

10 (a) Explain why A and A^{-1} have the same condition number.

  (b) Explain why A and A^{T} have the same norm, based on $\lambda(A^{\mathrm{T}}A)$ and $\lambda(AA^{\mathrm{T}})$.

11 Estimate the condition number of the ill-conditioned matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix}$.

12 Why is the determinant of A no good as a norm? Why is it no good as a condition number?

13   (Suggested by C. Moler and C. Van Loan.) Compute $b - Ay$ and $b - Az$ when

$$b = \begin{bmatrix} .217 \\ .254 \end{bmatrix} \quad A = \begin{bmatrix} .780 & .563 \\ .913 & .659 \end{bmatrix} \quad y = \begin{bmatrix} .341 \\ -.087 \end{bmatrix} \quad z = \begin{bmatrix} .999 \\ -1.0 \end{bmatrix}.$$

Is y closer than z to solving $Ax = b$? Answer in two ways: Compare the *residual* $b - Ay$ to $b - Az$. Then compare y and z to the true $x = (1, -1)$. Both answers can be right. Sometimes we want a small residual, sometimes a small Δx.

14   (a) Compute the determinant of A in Problem 13. Compute A^{-1}.

      (b) If possible compute $\|A\|$ and $\|A^{-1}\|$ and show that $c > 10^6$.

Problems 15–19 are about vector norms other than the usual $\|x\| = \sqrt{x \cdot x}$.

15   The "ℓ^1 norm" and the "ℓ^∞ norm" of $x = (x_1, \ldots, x_n)$ are

$$\|x\|_1 = |x_1| + \cdots + |x_n| \quad \text{and} \quad \|x\|_\infty = \max_{1 \le i \le n} |x_i|.$$

Compute the norms $\|x\|$ and $\|x\|_1$ and $\|x\|_\infty$ of these two vectors in $\mathbf{R}^5$:

$$x = (1, 1, 1, 1, 1) \qquad x = (.1, .7, .3, .4, .5).$$

16   Prove that $\|x\|_\infty \le \|x\| \le \|x\|_1$. Show from the Schwarz inequality that the ratios $\|x\|/\|x\|_\infty$ and $\|x\|_1/\|x\|$ are never larger than $\sqrt{n}$. Which vector $(x_1, \ldots, x_n)$ gives ratios equal to $\sqrt{n}$?

17   All vector norms must satisfy the *triangle inequality*. Prove that

$$\|x + y\|_\infty \le \|x\|_\infty + \|y\|_\infty \quad \text{and} \quad \|x + y\|_1 \le \|x\|_1 + \|y\|_1.$$

18   Vector norms must also satisfy $\|cx\| = |c| \|x\|$. The norm must be positive except when $x = 0$. Which of these are norms for vectors (x_1, x_2) in $\mathbf{R}^2$?

$$\|x\|_A = |x_1| + 2|x_2| \qquad \|x\|_B = \min(|x_1|, |x_2|)$$

$$\|x\|_C = \|x\| + \|x\|_\infty \qquad \|x\|_D = \|Ax\| \quad \text{(this answer depends on } A).$$

Challenge Problems

19   Show that $x^\mathsf{T} y \le \|x\|_1 \|y\|_\infty$ by choosing components $y_i = \pm 1$ to make $x^\mathsf{T} y$ as large as possible.

20   The eigenvalues of the $-1, 2, -1$ difference matrix K are $\lambda = 2 - 2\cos(j\pi/n{+}1)$. Estimate $\lambda_{\min}$ and $\lambda_{\max}$ and $c = \mathbf{cond}(K) = \lambda_{\max}/\lambda_{\min}$ as n increases: $c \approx Cn^2$ with what constant C?

      Test this estimate with $\mathbf{eig}(K)$ and $\mathbf{cond}(K)$ for $n = 10, 100, 1000$.