11.2 Norms and Condition Numbers

How do we measure the size of a matrix? For a vector, the length is $\|\boldsymbol{x}\|$. 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 $\|\boldsymbol{x}+\boldsymbol{y}\|$ is not greater than $\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$. The length of $2\boldsymbol{x}$ or $-2\boldsymbol{x}$ is doubled to $2\|\boldsymbol{x}\|$. The same rules will apply to matrix norms:

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

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

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

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

The norm of A ***is the largest ratio*** $\|A\boldsymbol{x}\|/\|\boldsymbol{x}\|$:
$$\|A\| = \max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}. \tag{3}$$

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

Example 1 If A is the identity matrix I, the ratios are $\|\boldsymbol{x}\|/\|\boldsymbol{x}\|$. Therefore $\|I\|=1$. If A is an orthogonal matrix Q, lengths are again preserved: $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$. 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):

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

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

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

Choose $\boldsymbol{x}$ to be the eigenvector with maximum eigenvalue. Then $\|A\boldsymbol{x}\|/\|\boldsymbol{x}\|$ equals $\lambda_{\max}$. The point is that no other $\boldsymbol{x}$ 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 $\|\Lambda\boldsymbol{x}\|/\|\boldsymbol{x}\|$. 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 $\|A\boldsymbol{x}\| = \|\lambda \boldsymbol{x}\| = |\lambda|$ times $\|\boldsymbol{x}\|$. The $\boldsymbol{x}$ 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:

$$\|\boldsymbol{A}\| > \boldsymbol{\lambda}_{\max} \quad A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{has norm} \quad \|A\| = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = 2.$$

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

It is the ***symmetric matrix*** $A^{\mathrm{T}}A$, not the unsymmetric A, that has eigenvector $\boldsymbol{x} = (0, 1)$. 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*** $\boldsymbol{\lambda}_{\max}(A^{\mathrm{T}}A)$***:***

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

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

No $\boldsymbol{x}$ can give a larger ratio. The symmetric matrix $A^{\mathrm{T}}A$ has eigenvalues $\lambda_1, \ldots, \lambda_n$ and orthonormal eigenvectors $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n$. Every $\boldsymbol{x}$ is a combination of those vectors. Try this combination in the ratio and remember that $\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_j = 0$:

$$\frac{\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}} = \frac{(c_1\boldsymbol{q}_1 + \cdots + c_n\boldsymbol{q}_n)^{\mathrm{T}}(c_1\lambda_1\boldsymbol{q}_1 + \cdots + c_n\lambda_n\boldsymbol{q}_n)}{(c_1\boldsymbol{q}_1 + \cdots + c_n\boldsymbol{q}_n)^{\mathrm{T}}(c_1\boldsymbol{q}_1 + \cdots + c_n\boldsymbol{q}_n)} = \frac{c_1^2\lambda_1 + \cdots + c_n^2\lambda_n}{c_1^2 + \cdots + c_n^2}.$$

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

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

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 $\boldsymbol{x}$.

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

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

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

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 $\Delta\boldsymbol{x}$ will be divided by 1000, but so will the exact solution $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. The ***relative error*** $\|\Delta\boldsymbol{x}\|/\|\boldsymbol{x}\|$ will stay the same. It is this relative change in $\boldsymbol{x}$ that should be compared to the relative change in $\boldsymbol{b}$.

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 $A\boldsymbol{x} = \boldsymbol{b}$.

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

Condition number *c*
$$\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \le c\frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}. \tag{6}$$

If the problem error is ΔA (error in A instead of $\boldsymbol{b}$), still c controls $\Delta\boldsymbol{x}$:

Error ΔA in A
$$\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x} + \Delta\boldsymbol{x}\|} \le c\frac{\|\Delta A\|}{\|A\|}. \tag{7}$$

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

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

Multiply the left sides to get $\|\boldsymbol{b}\|\,\|\Delta\boldsymbol{x}\|$, and multiply the right sides to get $c\|\boldsymbol{x}\|\,\|\Delta\boldsymbol{b}\|$. Divide both sides by $\|\boldsymbol{b}\|\,\|\boldsymbol{x}\|$. The left side is now the relative error $\|\Delta\boldsymbol{x}\|/\|\boldsymbol{x}\|$. 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 $\Delta\boldsymbol{b}$ in the error equation:

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

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

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

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

The error $\Delta\boldsymbol{b}$ depends on computer roundoff and on the original measurements of $\boldsymbol{b}$. 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 $\Delta\boldsymbol{x}$ can be unacceptable.

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

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

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 $\boldsymbol{x}$ small, choose $\boldsymbol{b}$ along the first eigenvector $(1, 0)$. To make $\Delta\boldsymbol{x}$ large, choose $\Delta\boldsymbol{b}$ along the second eigenvector $(0, 1)$. Then $\boldsymbol{x} = \frac{1}{6}\boldsymbol{b}$ and $\Delta\boldsymbol{x} = \frac{1}{2}\Delta\boldsymbol{b}$. The ratio $\|\Delta\boldsymbol{x}\|/\|\boldsymbol{x}\|$ is exactly $c = 3$ times the ratio $\|\Delta\boldsymbol{b}\|/\|\boldsymbol{b}\|$.

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 $\|B\|$:

$$\|AB\boldsymbol{x}\| \le \|A\|\,\|B\boldsymbol{x}\| \le \|A\|\,\|B\|\,\|\boldsymbol{x}\|.$$

From the ratio of $\|AB\boldsymbol{x}\|$ to $\|\boldsymbol{x}\|$, 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 $\|(A+B)\boldsymbol{x}\| \le \|A\boldsymbol{x}\| + \|B\boldsymbol{x}\|$ for every $\boldsymbol{x}$?

(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 $A\boldsymbol{x} = \lambda\boldsymbol{x}$.

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 $\boldsymbol{b} - A\boldsymbol{y}$ and $\boldsymbol{b} - A\boldsymbol{z}$ when

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

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

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 $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}$.

15 The "$\boldsymbol{\ell}^1$ norm" and the "$\boldsymbol{\ell}^\infty$ norm" of $\boldsymbol{x} = (x_1, \ldots, x_n)$ are

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

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

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

16 Prove that $\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\| \le \|\boldsymbol{x}\|_1$. Show from the Schwarz inequality that the ratios $\|\boldsymbol{x}\|/\|\boldsymbol{x}\|_\infty$ and $\|\boldsymbol{x}\|_1/\|\boldsymbol{x}\|$ 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

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

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

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

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

Challenge Problems

19 Show that $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} \le \|\boldsymbol{x}\|_1\,\|\boldsymbol{y}\|_\infty$ by choosing components $y_i = \pm 1$ to make $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$ 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 **eig**(K) and **cond**(K) for $n = 10, 100, 1000$.