Starting with Two Matrices

Gilbert Strang, MIT

1. INTRODUCTION The first sections of this paper represent an imaginary lecture, very near the beginning of a linear algebra course. We chose two matrices A and C, on the principle that examples are amazingly powerful. The reader is requested to be exceptionally patient, suspending all prior experience—and suspending also any hunger for precision and proof. Please allow a partial understanding to be established first.

I believe there is value in *naming* these matrices. The words "difference matrix" and "sum matrix" tell how they act. It is the action of matrices, when we form Ax and Cx and Sb, that makes linear algebra such a dynamic and beautiful subject.

2. A FIRST EXAMPLE In the future I will begin my linear algebra class with these three vectors a_1, a_2, a_3:

$$a_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad a_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \qquad a_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The fundamental operation on vectors is to take *linear combinations*. Multiply these vectors a_1, a_2, a_3 by numbers x_1, x_2, x_3 and add. This produces the linear combination $x_1a_1 + x_2a_2 + x_3a_3 = b$:

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{1}$$

(I am omitting words that would be spoken while writing that vector equation.) A key step is to rewrite (1) as a matrix equation:

Put the vectors a_1, a_2, a_3 into the columns of a matrix A
Put the multipliers x_1, x_2, x_3 into a vector x

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Then A times x is exactly $x_1a_1+x_2a_2+x_3a_3$. This is more than a definition of Ax, because the rewriting brings a crucial change in viewpoint. At first, the x's were multiplying the a's. Now, the matrix A is multiplying the vector x. **The matrix acts on** x, to give a combination of the columns of A:

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{2}$$

When the x's are known, the matrix A takes their differences. We could imagine an unwritten $x_0 = 0$, and put in $x_1 - x_0$ to complete the pattern. A is a **difference matrix**.

One more step to a new viewpoint. Suppose the x's are not known but the b's are known. Then $Ax = b$ becomes an equation for x, not an equation for b. We start with the differences (the b's) and ask which x's have those differences.

Linear algebra is always interested first in $b = 0$:

$$Ax = 0 \quad \text{is} \quad \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \qquad \text{Then} \quad \begin{matrix} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \end{matrix} \tag{3}$$

For this matrix, the only solution to $Ax = 0$ is $x = 0$. That may seem automatic but it's not. A key word in linear algebra (we are foreshadowing its importance) describes this situation. These column vectors a_1, a_2, a_3 are **independent**. The combination $x_1a_1 + x_2a_2 + x_3a_3$ is $Ax = 0$ only when all the x's are zero.

Move now to nonzero differences $b_1 = 1, b_2 = 3, b_3 = 5$. Is there a choice of x_1, x_2, x_3 that produces those differences? Solving the three equations in forward order, the x's are $1, 4, 9$:

$$\boldsymbol{Ax = b} \qquad \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \quad \text{leads to} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix}. \tag{4}$$

This case $x = 1, 4, 9$ has special interest. When the b's are the odd numbers in order, the x's are the perfect squares in order. But linear algebra is not number theory—forget that special case ! For any b_1, b_2, b_3 there is a neat formula for x_1, x_2, x_3:

$$\begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{leads to} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_1 + b_2 \\ b_1 + b_2 + b_3 \end{bmatrix}. \tag{5}$$

This general solution includes the examples with $b = 0, 0, 0$ (when $x = 0, 0, 0$) and $b = 1, 3, 5$ (when $x = 1, 4, 9$). One more insight will complete the example.

We started with a linear combination of a_1, a_2, a_3 to get b. Now b is given and equation (5) finds x. That solution shows three new vectors whose combination gives x:

$$x = b_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + b_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{6}$$

This is beautiful. The equation $Ax = b$ is solved by $x = Sb$. The matrix S is the "inverse" of the matrix A. The difference matrix is inverted by the **sum matrix**. Where A took differences of the x's, the new matrix S in (6) takes sums of the b's.

The linear algebra symbol for the inverse matrix is A^{-1} (not $1/A$). Thus $S = A^{-1}$ and also $A = S^{-1}$. This example shows how linear algebra goes in parallel with calculus. Sums are the inverse of differences, and integration S is the inverse of differentiation A:

$$Ax = \frac{dx}{dt} = b(t) \quad \text{is solved by} \quad x(t) = Sb = \int_0^t b. \tag{7}$$

The student who notices that the integral starts at $x(0) = 0$, and connects this to the earlier suggestion that $x_0 = 0$, is all too likely to become a mathematician.

3. THE SECOND EXAMPLE This example begins with almost the same three vectors—only one component is changed:

$$c_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad c_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \qquad c_3 = \begin{bmatrix} \mathbf{-1} \\ 0 \\ 1 \end{bmatrix}.$$

The combination $x_1c_1 + x_2c_2 + x_3c_3$ is again a matrix multiplication Cx:

$$Cx = \begin{bmatrix} & & \\ c_1 & c_2 & c_3 \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \quad (8)$$

With the new vector in the third column, C is a "cyclic" difference matrix. The differences of x's with "wraparound" give the new b's. The reverse direction begins with the b's and asks for the x's.

We always start with $0, 0, 0$ as the b's. You will see the change: nonzero x's can have zero differences. As long as the x's are equal, their differences will all be zero:

$$\begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{is solved by} \quad x = \begin{bmatrix} x_1 \\ x_1 \\ x_1 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \quad (9)$$

The zero solution $x = 0$ is included (when $x_1 = 0$). But $1, 1, 1$ and $2, 2, 2$ and π, π, π are also solutions—all these constant vectors have zero differences and solve $Cx = 0$. The columns c_1, c_2, c_3 are **dependent** and not independent.

This misfortune produces a new difficulty, when we try to solve $Cx = b$:

$$\begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{cannot be solved unless} \quad b_1 + b_2 + b_3 = 0.$$

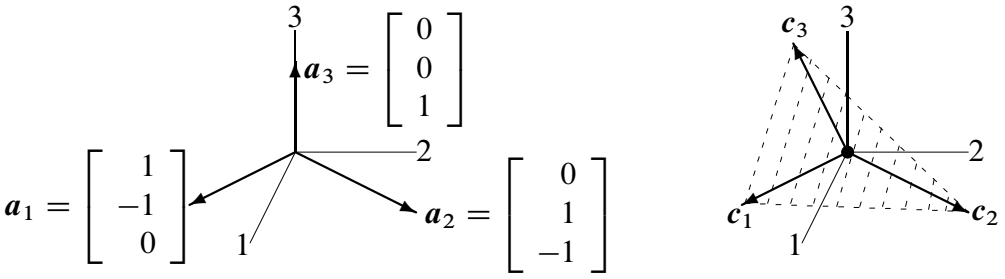
The three left sides add to zero, because x_3 is now cancelled by $-x_3$. So the b's on the right side must add to zero. There is no solution like equation (5) for every b_1, b_2, b_3. There is no inverse matrix like S. The cyclic matrix C is *singular*.

4. SUMMARY Both examples began by putting vectors into the columns of a matrix. Combinations of the columns (with multipliers x) became Ax and Cx. Difference matrices A and C (non-cyclic and cyclic) multiplied x—that was an important switch in thinking. The details of those column vectors made $Ax = b$ solvable for all b, while $Cx = b$ is not always solvable. The words that express the contrast between A and C are a crucial part of the language of linear algebra:

The vectors a_1, a_2, a_3 are independent.
The nullspace for $Ax = 0$ contains only $x = 0$.
The equation $Ax = b$ is solved by $x = Sb$.
The square matrix A has the inverse matrix $S = A^{-1}$.

The vectors c_1, c_2, c_3 are dependent.
The nullspace for $Cx = 0$ contains every "constant vector" x_1, x_1, x_1.
The equation $Cx = b$ cannot be solved unless $b_1 + b_2 + b_3 = 0$.
C has no inverse matrix.

A picture of the three vectors, the a's on the left and the c's on the right, explains the difference in a useful way. On the left, the three directions are *independent*. The a's don't lie in a plane. Their combinations $x_1a_1 + x_2a_2 + x_3a_3$ produce every three-dimensional vector b. The good multipliers x_1, x_2, x_3 are given by $x = Sb$.



On the right, the vectors c_1, c_2, c_3 are *dependent*. The c's do lie in a plane. Each vector has components adding to $1 - 1 = 0$, so all combinations of these vectors will have $b_1 + b_2 + b_3 = 0$. The differences $x_1 - x_3, x_2 - x_1, x_3 - x_2$ can never be $1, 1, 1$.

Almost unconsciously, those examples illustrate one way of teaching mathematics. The ideas and the words are used before they are fully defined. I believe we learn our own language this way—by hearing words, trying to use them, making mistakes (matrices are not dependent, vectors are not singular), and eventually getting it right. A proper definition is certainly needed, it is not at all an afterthought. But it could be an afterword.

5. FUTURE LECTURE: GRAM-SCHMIDT PRODUCES ORTHOGONAL VECTORS Sections 5, 6, 7 are not for the student, at least not yet. The reader deserves some fun too. We look at the same matrices A and C with three new questions—orthogonalization, eigenvalues, and singular values.

The vectors a_1, a_2, a_3 are not orthogonal. Then the matrix $A^{\mathrm{T}}A$ is not diagonal. Since that symmetric matrix appears over and over in applications, it is safer to orthogonalize the vectors in advance. Gram-Schmidt is the easiest way to do it (but not the best way [1, 3]). By removing the projections of a_2 onto a_1 and a_3 onto a_2, we go from the a's to orthogonal vectors:

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ are orthogonalized to } \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \\ -1 \end{bmatrix} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \qquad (10)$$

When those are normalized to unit vectors, they are the columns of an orthogonal matrix Q. It is connected to the original A by $A = QR$, where R is upper triangular (because all steps involved earlier vectors and not later vectors [4]). The real pleasure is the pattern of fractions in equation (10), leading to a neat orthogonal basis in every dimension n.

And the real content, when a future lecture orthogonalizes the second example c_1, c_2, c_3, is to see the Gram-Schmidt process *fall short*:

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ are orthogonalized to } \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \\ -1 \end{bmatrix} \begin{bmatrix} \\ \\ \end{bmatrix}. \tag{11}$$

The third column of Q is empty because c_1, c_2, c_3 lie in a plane.

6. EIGENVALUES AND EIGENVECTORS The eigenvectors of the triangular matrix A are its diagonal entries $1, 1, 1$. The eigenvectors (or lack of them) are more interesting: the only eigenvectors are multiples of the last column $0, 0, 1$. The Jordan form J is a single 3 by 3 block, reached by a matrix G with "generalized eigenvectors":

$$J = G^{-1}AG = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \text{ with } G = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & -1 \\ 1 & 1 & 0 \end{bmatrix}. \tag{12}$$

This gives an example for discussion, but it's not at the heart of the course. It makes a sharp contrast with C.

In the matrix C, each column is a cyclic permutation P times the previous column. Then $C = I - P$ is a **circulant matrix** [1], and its diagonals wrap around (unlike A). Circulants are polynomials in the cyclic permutation matrix P, so they all commute.

The eigenvectors of every circulant (including P and C) are the columns of the Fourier matrix F. All entries in F are nth roots of 1; our example has $n = 3$. The cube root $w = e^{i\theta}$ has $\theta = 2\pi/3$. Then $w^3 = 1$ is the key to verifying the three eigenvectors of P in the columns of F:

$$PF = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & w & w^2 \\ 1 & w^2 & w^4 \end{bmatrix} = \begin{bmatrix} 1 & w^2 & w^4 \\ 1 & w^3 & w^6 \\ 1 & w^4 & w^8 \end{bmatrix}. \tag{13}$$

Each column of (13) shows $Px = \lambda x$. The columns of F are multiplied by the eigenvalues $1, w^2, w^4$ of P. Then the eigenvalues of $C = I - P$ must be $1-1, 1-w^2, 1-w^4$. That eigenvalue $1 - 1 = 0$ confirms that C is singular. The other two eigenvalues are complex.

Circulant matrices play a major role as "filters" in signal processing. The output Cx is a convolution of the input x with the column c_1. Again the matrix acts! Where $x_0 = 0$ was natural for the non-cyclic matrix A, $x_0 = x_3$ is natural for C. There is a small closed friendly world of periodic vectors and circulant matrices, in which $w^3 = 1$ (or $w^n = 1$) and everything repeats.

One more important point. Since circulants commute, $C^{\mathrm{T}}C$ equals CC^{T}. Then C is a **normal matrix**, and its eigenvectors in F are orthogonal. This is one reason for the great success of the Discrete Fourier Transform:

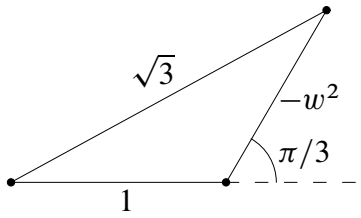
Orthogonal eigenvectors of C
$$\overline{F}^{\mathrm{T}}F = \begin{bmatrix} 3 & & \\ & 3 & \\ & & 3 \end{bmatrix}. \tag{14}$$

7. SINGULAR VALUES AND SINGULAR VECTORS Every matrix A has orthonormal *singular vectors*. The equation $Ax_i = \lambda_i x_i$ changes to $Av_i = \sigma_i u_i$. There are *two* orthonormal sequences, the v's and the u's. If those are the columns of V and U, the vector equation $Av_i = \sigma_i u_i$ becomes a matrix equation $AV = U\Sigma$. Then $A = U\Sigma V^{-1}$ is the celebrated **Singular Value Decomposition** of A, with $V^{-1} = \overline{V}^{\mathrm{T}}$.

A personal note: Numerical linear algebra has just lost its foremost leader, Gene Golub of Stanford University. Gene will always be associated with the SVD. He showed us how to compute it and how to use it—this small paper is dedicated to a most remarkable friend.

The SVD separates A into r rank one pieces $\sigma_i u_i \overline{v}_i^{\mathrm{T}}$ *in order of their importance* (based on the sizes of the $\sigma_i > 0$). When the rank r of A is less than m or n, there will be $m - r$ more u's and $n - r$ more v's (orthonormal bases for the nullspaces of A^{T} and A).

Naturally we hope that the difference matrices A and C have attractive decompositions. The eigenvectors of C are already orthogonal (in F), and normalizing to unit vectors v_i gives $V = F/\sqrt{3}$. But the eigenvalues $1 - w^2$ and $1 - w^4$ are not the σ's, real and positive. We need to adjust $Cv = \lambda v$ to $Cv = \sigma u$, by writing each complex λ as $\sigma e^{i\theta}$ (and then u is $e^{i\theta} v$). The particular matrix C has $\sigma = |\lambda| = \sqrt{3}$ and $\theta = \pm\pi/6$, because the two terms in $\lambda = 1 - w^2$ and $1 - w^4$ form 30-30-120 isosceles triangles.



shows $\lambda = \sigma e^{i\theta} = \sqrt{3} e^{i\pi/6}$

The decomposition of the first example A into $U\Sigma V^{\mathrm{T}}$ is especially interesting. This difference matrix is non-cyclic and non-normal. Entirely different bases of u's and v's are required to reach $Av_i = \sigma_i u_i$. The rules say that the u's and v's are eigenvectors of AA^{T} and $A^{\mathrm{T}}A$. (They have the same positive eigenvalues σ_i^2.) Our example leads to second difference matrices:

$$AA^{\mathrm{T}} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \qquad A^{\mathrm{T}}A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \tag{15}$$

Since the eigenfunctions in $-d^2u/dx^2 = \lambda u$ are sines and cosines, we may hope that the eigenvectors here are *discrete* sines and cosines. Wonderfully, they are.

That first row of AA^{T} suggests a boundary condition $du/dx = 0$. Then the u's will be cosines. The first row of $A^{\mathrm{T}}A$ has already been linked to $v_0 = 0$. Then the v's will be sines. The SVD of the difference matrix A has a meaning for *discrete calculus*:

$$Av = \sigma u \quad \text{is a discrete form of} \quad \frac{d}{dx} \sin \sigma x = \sigma \cos \sigma x. \tag{16}$$

With patience and MATLAB or *Mathematica*, these sines and cosines in V and U can be made to appear:

$$V = \begin{bmatrix} \sin\frac{\pi}{7} & \sin\frac{3\pi}{7} & \sin\frac{5\pi}{7} \\ \sin\frac{2\pi}{7} & \sin\frac{6\pi}{7} & \sin\frac{10\pi}{7} \\ \sin\frac{3\pi}{7} & \sin\frac{9\pi}{7} & \sin\frac{15\pi}{7} \end{bmatrix} \qquad \begin{aligned} \sigma_1 &= 2\sin\frac{\pi}{14} \\ \sigma_2 &= 2\sin\frac{3\pi}{14} \\ \sigma_3 &= 2\sin\frac{5\pi}{14} \end{aligned}$$

$$U = \begin{bmatrix} \cos\frac{\pi}{14} & \cos\frac{3\pi}{14} & \cos\frac{5\pi}{14} \\ \cos\frac{3\pi}{14} & \cos\frac{9\pi}{14} & \cos\frac{15\pi}{14} \\ \cos\frac{5\pi}{14} & \cos\frac{15\pi}{14} & \cos\frac{25\pi}{14} \end{bmatrix} \qquad \begin{aligned} Av_1 &= \sigma_1 u_i \\ Av_2 &= \sigma_2 u_2 \\ Av_3 &= \sigma_3 u_3 \end{aligned}$$

The rows of U are the rows of V in reverse order, because the same is true in (15). All columns have the *same three numbers* (apart from signs). It is remarkable to find orthogonal matrices with this property. I don't know a full list of such matrices.

8. SECOND DIFFERENCE MATRICES We are close to something important for calculus and differential equations and linear algebra. Allow me to mention it here, after the main point (the early lecture) is completed. That lecture chose "backward differences" but there were really three good options:

Forward difference	$\dfrac{u(x+h)-u(x)}{h}$
Backward difference	$\dfrac{u(x)-u(x-h)}{h}$
Centered difference	$\dfrac{u(x+h)-u(x-h)}{2h}$

When $u(x) = x$, all three give the correct derivative $u' = 1$. But for $u(x) = x^2$, only the centered difference gives $2x$. It has second order accuracy, and is preferred in scientific computation—when a differential equation becomes discrete, which is calculus in reverse.

One more big step. A forward difference of a backward difference is a second difference. The numbers $1, -1$ go up a level to $1, -2, 1$:

Second difference	$\dfrac{u(x+h)-2u(x)+u(x-h)}{h^2}$

This is $u'' = d/dx(du/dx)$, made discrete and centered. It is the start of computational science, and we have to express it with matrices. That step starts from n values $u_1, u_2, \ldots, u_n$ (take $u_0 = 0$ and $u_{n+1} = 0$) and produces n second differences. Go carefully

here, the student is learning to convert formulas at n points into a matrix multiplication (and division by h^2):

$$\frac{d^2u}{dx^2} \quad \text{matches} \quad \frac{-Ku}{h^2} \quad \text{with} \quad Ku = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}. \tag{17}$$

I reversed signs to $-1, 2, -1$ to reach my favorite matrix K. It is a "second difference matrix" and it has wonderful properties. May I beg you to write this matrix K on the board, along with the analogous 4 by 4 matrix, and *ask the class for their properties*:

First answer always: K is symmetric
Second answer frequently: K has three nonzero diagonals (tridiagonal)
Third answer sometimes: K is invertible (det $K_n = n + 1$)
Fourth answer never: K is positive definite

But that fourth answer—the idea of a *positive definite matrix*—unifies the whole course. K is seen and named much earlier. Its pivots are positive and its eigenvalues are positive. It appears as $A^{\mathrm{T}}A$ when the backward difference matrix A has a fourth row $0, 0, -1$. At that point we know about rectangular matrices and least squares, which is their number one application and invariably leads to $A^{\mathrm{T}}A$.

Could I end where I began, with the essential words of linear algebra? A is now rectangular, and the idea of *rank* has to be learned—by example, by definition, and by the key theorems $\mathrm{rank}(A) = \mathrm{rank}(A^{\mathrm{T}}) = \mathrm{rank}(A^{\mathrm{T}}A)$:

Vectors are independent $\quad$ A has full column rank $\quad$ $A^{\mathrm{T}}A$ is invertible

$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \qquad \begin{pmatrix} \text{and symmetric} \\ \text{positive definite} \end{pmatrix}$$

REFERENCES

1. Philip Davis, *Circulant Matrices*, Wiley, 1979.

2. Gene Golub and Charles Van Loan, *Matrix Computations*, Johns-Hopkins University Press, 1996.

3. Gilbert Strang, *Computational Science and Engineering*, Wellesley-Cambridge Press, 2007.

4. Gilbert Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, 2003.