INTRODUCTION
TO
LINEAR ALGEBRA

Third Edition

GILBERT STRANG
Massachusetts Institute of Technology

WELLESLEY-CAMBRIDGE PRESS
Box 812060 Wellesley MA 02482

TABLE OF CONTENTS

For square matrices, an inverse on one side is automatically an inverse on the other side. If $AB = I$ then automatically $BA = I$. In that case B is A^{-1}. This is very useful to know but we are not ready to prove it.

Example 3 Suppose F subtracts 4 times row 2 from row 3, and F^{-1} adds it back:

$$F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \quad \text{and} \quad F^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}.$$

Now multiply F by the matrix E in Example 2 to find FE. Also multiply E^{-1} times F^{-1} to find $(FE)^{-1}$. Notice the orders FE and $E^{-1}F^{-1}$!

$$FE = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 20 & -4 & 1 \end{bmatrix} \quad \text{is inverted by} \quad E^{-1}F^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}. \tag{6}$$

The result is strange but correct. The product FE contains "20" but its inverse doesn't. E subtracts 5 times row 1 from row 2. Then F subtracts 4 times the *new* row 2 (changed by row 1) from row 3. ***In this order FE, row 3 feels an effect from row 1.***

In the order $E^{-1}F^{-1}$, that effect does not happen. First F^{-1} adds 4 times row 2 to row 3. After that, E^{-1} adds 5 times row 1 to row 2. There is no 20, because row 3 doesn't change again. ***In this order, row 3 feels no effect from row 1.***

For elimination with normal order FE, ***the product of inverses $E^{-1}F^{-1}$ is quick. The multipliers fall into place below the diagonal of 1's.***

This special property of $E^{-1}F^{-1}$ and $E^{-1}F^{-1}G^{-1}$ will be useful in the next section. We will explain it again, more completely. In this section our job is A^{-1}, and we expect some serious work to compute it. Here is a way to organize that computation.

Calculating A^{-1} by Gauss-Jordan Elimination

I hinted that A^{-1} might not be explicitly needed. The equation $Ax = b$ is solved by $x = A^{-1}b$. But it is not necessary or efficient to compute A^{-1} and multiply it times b. *Elimination goes directly to x.* Elimination is also the way to calculate A^{-1}, as we now show. The Gauss-Jordan idea is to solve $AA^{-1} = I$, *finding each column of A^{-1}.*

A multiplies the first column of A^{-1} (call that x_1) to give the first column of I (call that e_1). This is our equation $Ax_1 = e_1 = (1, 0, 0)$. Each of the columns x_1, x_2, x_3 of A^{-1} is multiplied by A to produce a column of I:

$$AA^{-1} = A\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} = I. \tag{7}$$

To invert a 3 by 3 matrix A, we have to solve three systems of equations: $Ax_1 = e_1$ and $Ax_2 = e_2 = (0, 1, 0)$ and $Ax_3 = e_3 = (0, 0, 1)$. This already shows why computing A^{-1} is expensive. We must solve n equations for its n columns. To solve $Ax = b$ without A^{-1}, we deal only with *one* column.

In defense of A^{-1}, we want to say that its cost is not n times the cost of one system $Ax = b$. Surprisingly, the cost for n columns is only multiplied by 3. This saving is because the n equations $Ax_i = e_i$ all involve the same matrix A. Working with the right sides is relatively cheap, because elimination only has to be done once on A. The complete A^{-1} needs n^3 elimination steps, where a single x needs $n^3/3$. The next section calculates these costs.

The ***Gauss-Jordan method*** computes A^{-1} by solving all n equations together. Usually the "augmented matrix" has one extra column b, from the right side of the equations. Now we have three right sides e_1, e_2, e_3 (when A is 3 by 3). They are the columns of I, so the augmented matrix is really the block matrix $[\,A \quad I\,]$. Here is a worked-out example when A has 2's on the main diagonal and -1's next to the 2's:

$$\begin{bmatrix} A & e_1 & e_2 & e_3 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{bmatrix} \quad \textbf{Start Gauss-Jordan}$$

$$\to \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ \mathbf{0} & \mathbf{\tfrac{3}{2}} & \mathbf{-1} & \mathbf{\tfrac{1}{2}} & \mathbf{1} & \mathbf{0} \\ 0 & -1 & 2 & 0 & 0 & 1 \end{bmatrix} \quad (\tfrac{1}{2}\ \textbf{\textit{row}}\ \mathbf{1} + \textbf{\textit{row}}\ \mathbf{2})$$

$$\to \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \tfrac{3}{2} & -1 & \tfrac{1}{2} & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{\tfrac{4}{3}} & \mathbf{\tfrac{1}{3}} & \mathbf{\tfrac{2}{3}} & \mathbf{1} \end{bmatrix} \quad (\tfrac{2}{3}\ \textbf{\textit{row}}\ \mathbf{2} + \textbf{\textit{row}}\ \mathbf{3})$$

We are now halfway. The matrix in the first three columns is U (upper triangular). The pivots $2, \frac{3}{2}, \frac{4}{3}$ are on its diagonal. Gauss would finish by back substitution. The contribution of Jordan is *to continue with elimination*! He goes all the way to the ***"reduced echelon form"***. Rows are added to rows above them, to produce ***zeros above the pivots***:

$$\to \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ \mathbf{0} & \mathbf{\tfrac{3}{2}} & \mathbf{0} & \mathbf{\tfrac{3}{4}} & \mathbf{\tfrac{3}{2}} & \mathbf{\tfrac{3}{4}} \\ 0 & 0 & \tfrac{4}{3} & \tfrac{1}{3} & \tfrac{2}{3} & 1 \end{bmatrix} \quad (\tfrac{3}{4}\ \textbf{\textit{row}}\ \mathbf{3} + \textbf{\textit{row}}\ \mathbf{2})$$

$$\to \begin{bmatrix} \mathbf{2} & \mathbf{0} & \mathbf{0} & \mathbf{\tfrac{3}{2}} & \mathbf{1} & \mathbf{\tfrac{1}{2}} \\ 0 & \tfrac{3}{2} & 0 & \tfrac{3}{4} & \tfrac{3}{2} & \tfrac{3}{4} \\ 0 & 0 & \tfrac{4}{3} & \tfrac{1}{3} & \tfrac{2}{3} & 1 \end{bmatrix} \quad (\tfrac{2}{3}\ \textbf{\textit{row}}\ \mathbf{2} + \textbf{\textit{row}}\ \mathbf{1})$$

The last Gauss-Jordan step is to divide each row by its pivot. The new pivots are 1. We have reached I in the first half of the matrix, because A is invertible. ***The three columns of A^{-1} are in the second half of*** $[\, I \;\, A^{-1}\,]$:

$$\begin{matrix}(\text{divide by } 2)\\ (\text{divide by } \frac{3}{2})\\ (\text{divide by } \frac{4}{3})\end{matrix} \quad \begin{bmatrix} 1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix} = \begin{bmatrix} I & x_1 & x_2 & x_3 \end{bmatrix}.$$

Starting from the 3 by 6 matrix $[\, A \;\; I\,]$, we ended with $[\, I \;\; A^{-1}\,]$. Here is the whole Gauss-Jordan process on one line:

$$\textbf{\textit{Multiply}} \;\; \begin{bmatrix} A & I \end{bmatrix} \;\; \textbf{\textit{by}} \;\; A^{-1} \;\; \textbf{\textit{to get}} \;\; [I \; A^{-1}].$$

The elimination steps gradually create the inverse matrix. For large matrices, we probably don't want A^{-1} at all. But for small matrices, it can be very worthwhile to know the inverse. We add three observations about this particular A^{-1} because it is an important example. We introduce the words *symmetric*, *tridiagonal*, and *determinant* (Chapter 5):

1. A is ***symmetric*** across its main diagonal. So is A^{-1}.

2. A is ***tridiagonal*** (only three nonzero diagonals). But A^{-1} is a full matrix with no zeros. That is another reason we don't often compute A^{-1}.

3. The product of pivots is $2(\frac{3}{2})(\frac{4}{3}) = 4$. This number 4 is the ***determinant*** of A.

$$\boldsymbol{A^{-1}} \textbf{\textit{ involves division by the determinant}} \qquad A^{-1} = \frac{1}{4}\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}. \tag{8}$$

Example 4 Find A^{-1} by Gauss-Jordan elimination starting from $A = \begin{bmatrix} 2 & 3 \\ 4 & 7 \end{bmatrix}$. There are two row operations and then a division to put 1's in the pivots:

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{3} & 1 & 0 \\ \mathbf{4} & \mathbf{7} & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{2} & \mathbf{3} & 1 & 0 \\ \mathbf{0} & \mathbf{1} & \mathbf{-2} & \mathbf{1} \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} \mathbf{2} & \mathbf{0} & \mathbf{7} & \mathbf{-3} \\ 0 & 1 & -2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & \frac{7}{2} & -\frac{3}{2} \\ 0 & 1 & -\mathbf{2} & \mathbf{1} \end{bmatrix} = \begin{bmatrix} I & A^{-1} \end{bmatrix}.$$

The reduced echelon form of $[\, A \;\; I\,]$ is $[\, I \;\; A^{-1}\,]$. This A^{-1} involves division by the determinant $2 \cdot 7 - 3 \cdot 4 = 2$. The code for $X =$ inverse(A) has three important lines!

```
I = eye (n, n);          % Define the identity matrix
R = rref ([A I]);        % Eliminate on the augmented matrix
X = R( : , n + 1 : n + n) % Pick A^-1 from the last n columns of R
```

A must be invertible, or elimination will not reduce it (in the left half of R) to I.

Singular versus Invertible

We come back to the central question. Which matrices have inverses? The start of this section proposed the pivot test: $\boldsymbol{A^{-1}}$ ***exists exactly when A has a full set of n pivots.*** (Row exchanges allowed.) Now we can prove that by Gauss-Jordan elimination:

1. With n pivots, elimination solves all the equations $A\boldsymbol{x}_i = \boldsymbol{e}_i$. The columns $\boldsymbol{x}_i$ go into A^{-1}. Then $AA^{-1} = I$ and A^{-1} is at least a ***right-inverse***.

2. Elimination is really a sequence of multiplications by E's and P's and D^{-1}:

$$(D^{-1} \cdots E \cdots P \cdots E)A = I. \tag{9}$$

D^{-1} divides by the pivots. The matrices E produce zeros below and above the pivots. P will exchange rows if needed (see Section 2.7). The product matrix in equation (9) is evidently a ***left-inverse***. With n pivots we reach $A^{-1}A = I$.

The right-inverse equals the left-inverse. That was Note 2 in this section. So a square matrix with a full set of pivots will always have a two-sided inverse.

Reasoning in reverse will now show that A must have n pivots if $AC = I$. Then we deduce that C is also a left-inverse. Here is one route to those conclusions:

1. If A doesn't have n pivots, elimination will lead to a *zero row*.
2. Those elimination steps are taken by an invertible M. *So a row of MA is zero.*
3. If $AC = I$ then $MAC = M$. The zero row of MA, times C, gives a zero row of M.
4. The invertible matrix M can't have a zero row! A *must* have n pivots if $AC = I$.
5. Then equation (9) displays the left inverse in $BA = I$, and Note 2 proves $B = C$.

That argument took five steps, but the outcome is short and important.

2I A complete test for invertibility of a square matrix A comes from elimination. $\boldsymbol{A^{-1}}$ ***exists (and Gauss-Jordan finds it) exactly when A has n pivots.*** The full argument shows more:

$$\text{If} \quad AC = I \quad \text{then} \quad CA = I \quad \text{and} \quad C = A^{-1} \quad !$$

Example 5 If L is lower triangular with 1's on the diagonal, so is L^{-1}.

Use the Gauss-Jordan method to construct L^{-1}. Start by subtracting multiples of pivot rows from rows *below*. Normally this gets us halfway to the inverse, but for L it gets us all the way. L^{-1} appears on the right when I appears on the left:

$$\begin{bmatrix} L & I \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & 1 & 0 & 0 \\ \mathbf{3} & \mathbf{1} & \mathbf{0} & 0 & 1 & 0 \\ \mathbf{4} & \mathbf{5} & \mathbf{1} & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{matrix} \\ \to \\ \to \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -3 & 1 & 0 \\ 0 & 5 & 1 & -4 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \\ \text{(3 times row 1 from row 2)} \\ \text{(4 times row 1 from row 3)} \end{matrix}$$

$$\begin{matrix} \\ \\ \to \end{matrix} \begin{bmatrix} 1 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ 0 & 1 & 0 & \mathbf{-3} & \mathbf{1} & \mathbf{0} \\ 0 & 0 & 1 & \mathbf{11} & \mathbf{-5} & \mathbf{1} \end{bmatrix} = \begin{bmatrix} I & \mathbf{L^{-1}} \end{bmatrix}.$$

When L goes to I by elimination, I goes to L^{-1}. In other words, the product of elimination matrices $E_{32}E_{31}E_{21}$ is L^{-1}. All pivots are 1's (a full set). L^{-1} is lower triangular. The strange entry "11" in L^{-1} does not appear in $E_{21}^{-1}E_{31}^{-1}E_{32}^{-1} = L$.

■ REVIEW OF THE KEY IDEAS ■

1. The inverse matrix gives $AA^{-1} = I$ and $A^{-1}A = I$.

2. A is invertible if and only if it has n pivots (row exchanges allowed).

3. If $A\boldsymbol{x} = \mathbf{0}$ for a nonzero vector $\boldsymbol{x}$, then A has no inverse.

4. The inverse of AB is the reverse product $B^{-1}A^{-1}$.

5. The Gauss-Jordan method solves $AA^{-1} = I$ to find the n columns of A^{-1}. The augmented matrix $\begin{bmatrix} A & I \end{bmatrix}$ is row-reduced to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$.

■ WORKED EXAMPLES ■

2.5 A Three of these matrices are invertible, and three are singular. Find the inverse when it exists. Give reasons for noninvertibility (zero determinant, too few pivots, nonzero solution to $A\boldsymbol{x} = \mathbf{0}$) for the other three, in that order. The matrices A, B, C, D, E, F are

$$\begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix} \quad \begin{bmatrix} 4 & 3 \\ 8 & 7 \end{bmatrix} \quad \begin{bmatrix} 6 & 6 \\ 6 & 0 \end{bmatrix} \quad \begin{bmatrix} 6 & 6 \\ 6 & 6 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Solution

$$B^{-1} = \frac{1}{4}\begin{bmatrix} 7 & -3 \\ -8 & 4 \end{bmatrix} \quad C^{-1} = \frac{1}{36}\begin{bmatrix} 0 & 6 \\ 6 & -6 \end{bmatrix} \quad E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

A is not invertible because its determinant is $4 \cdot 6 - 3 \cdot 8 = 24 - 24 = 0$. D is not invertible because there is only one pivot; the second row becomes zero when the first row is subtracted. F is not invertible because a combination of the columns (the second column minus the first column) is zero—in other words $F\boldsymbol{x} = \mathbf{0}$ has the solution $\boldsymbol{x} = (-1, 1, 0)$.

Of course all three reasons for noninvertibility would apply to each of A, D, F.

2.5 B Apply the Gauss-Jordan method to find the inverse of this triangular "Pascal matrix" $A =$ abs(pascal(4,1)). You see **Pascal's triangle**—adding each entry to the entry on its left gives the entry below. The entries are "binomial coefficients":

$$\textbf{Triangular Pascal matrix} \quad A = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{1} & 0 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{1} & 0 \\ \mathbf{1} & \mathbf{3} & \mathbf{3} & \mathbf{1} \end{bmatrix}.$$

Solution Gauss-Jordan starts with $[A \;\; I]$ and produces zeros by subtracting row 1:

$$[A \;\; I] = \left[\begin{array}{cccc|cccc} \mathbf{1} & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{1} & 0 & 0 & 0 & 1 & 0 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{1} & 0 & 0 & 0 & 1 & 0 \\ \mathbf{1} & \mathbf{3} & \mathbf{3} & \mathbf{1} & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 & -1 & 0 & 0 & 1 \end{array}\right].$$

The next stage creates zeros below the second pivot, using multipliers 2 and 3. Then the last stage subtracts 3 times the new row 3 from the new row 4:

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 3 & 1 & 2 & -3 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \mathbf{-1} & \mathbf{1} & 0 & 0 \\ 0 & 0 & 1 & 0 & \mathbf{1} & \mathbf{-2} & \mathbf{1} & 0 \\ 0 & 0 & 0 & 1 & \mathbf{-1} & \mathbf{3} & \mathbf{-3} & \mathbf{1} \end{array}\right] = [I \;\; A^{-1}].$$

All the pivots were 1! So we didn't need to divide rows by pivots to get I. The inverse matrix A^{-1} looks like A itself, except odd-numbered diagonals are multiplied by -1.

Please notice that 4 by 4 matrix A^{-1}, we will see Pascal matrices again. The same pattern continues to n by n Pascal matrices—the inverse has "alternating diagonals".

Problem Set 2.5

1 Find the inverses (directly or from the 2 by 2 formula) of A, B, C:

$$A = \begin{bmatrix} 0 & 3 \\ 4 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 0 \\ 4 & 2 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 3 & 4 \\ 5 & 7 \end{bmatrix}.$$

3

VECTOR SPACES AND SUBSPACES

SPACES OF VECTORS ■ 3.1

To a newcomer, matrix calculations involve a lot of numbers. To you, they involve vectors. The columns of Ax and AB are linear combinations of n vectors—the columns of A. This chapter moves from numbers and vectors to a third level of understanding (the highest level). Instead of individual columns, we look at "spaces" of vectors. Without seeing ***vector spaces*** and especially their ***subspaces***, you haven't understood everything about $Ax = b$.

Since this chapter goes a little deeper, it may seem a little harder. That is natural. We are looking inside the calculations, to find the mathematics. The author's job is to make it clear. These pages go to the heart of linear algebra.

We begin with the most important vector spaces. They are denoted by $\mathbf{R}^1$, $\mathbf{R}^2$, $\mathbf{R}^3$, $\mathbf{R}^4$, ... Each space $\mathbf{R}^n$ consists of a whole collection of vectors. $\mathbf{R}^5$ contains all column vectors with five components. This is called "5-dimensional space."

DEFINITION ***The space $\mathbf{R}^n$ consists of all column vectors v with n components.***

The components of v are real numbers, which is the reason for the letter $\mathbf{R}$. A vector whose n components are complex numbers lies in the space $\mathbf{C}^n$.

The vector space $\mathbf{R}^2$ is represented by the usual xy plane. Each vector v in $\mathbf{R}^2$ has two components. The word "space", asks us to think of all those vectors—the whole plane. Each vector gives the x and y coordinates of a point in the plane.

Similarly the vectors in $\mathbf{R}^3$ correspond to points (x, y, z) in three-dimensional space. The one-dimensional space $\mathbf{R}^1$ is a line (like the x axis). As before, we print vectors as a column between brackets, or along a line using commas and parentheses:

$$\begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \text{ is in } \mathbf{R}^3, \quad (1, 1, 0, 1, 1) \text{ is in } \mathbf{R}^5, \quad \begin{bmatrix} 1+i \\ 1-i \end{bmatrix} \text{ is in } \mathbf{C}^2.$$

The great thing about linear algebra is that it deals easily with five-dimensional space. We don't draw the vectors, we just need the five numbers (or n numbers). To multiply $\boldsymbol{v}$ by 7, multiply every component by 7. Here 7 is a "scalar." To add vectors in $\mathbf{R}^5$, add them a component at a time. The two essential vector operations go on *inside the vector space*:

We can add any vectors in $\mathbf{R}^n$, and we can multiply any vector by any scalar.

"Inside the vector space" means that ***the result stays in the space.*** If $\boldsymbol{v}$ is the vector in $\mathbf{R}^4$ with components $1, 0, 0, 1$, then $2\boldsymbol{v}$ is the vector in $\mathbf{R}^4$ with components $2, 0, 0, 2$. (In this case 2 is the scalar.) A whole series of properties can be verified in $\mathbf{R}^n$. The commutative law is $\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}$; the distributive law is $c(\boldsymbol{v} + \boldsymbol{w}) = c\boldsymbol{v} + c\boldsymbol{w}$. There is a unique "zero vector" satisfying $\mathbf{0} + \boldsymbol{v} = \boldsymbol{v}$. Those are three of the eight conditions listed at the start of the problem set.

These eight conditions are required of every vector space. There are vectors other than column vectors, and vector spaces other than $\mathbf{R}^n$, and they have to obey the eight reasonable rules.

A real vector space is a set of "vectors" *together with rules for vector addition and for multiplication by real numbers*. The addition and the multiplication must produce vectors that are in the space. And the eight conditions must be satisfied (which is usually no problem). Here are three vector spaces other than $\mathbf{R}^n$:

M The vector space of ***all real*** 2 ***by*** 2 ***matrices.***
F The vector space of ***all real functions*** $f(x)$.
Z The vector space that consists only of a ***zero vector.***

In **M** the "vectors" are really matrices. In **F** the vectors are functions. In **Z** the only addition is $\mathbf{0} + \mathbf{0} = \mathbf{0}$. In each case we can add: matrices to matrices, functions to functions, zero vector to zero vector. We can multiply a matrix by 4 or a function by 4 or the zero vector by 4. The result is still in **M** or **F** or **Z**. The eight conditions are all easily checked.

The space **Z** is zero-dimensional (by any reasonable definition of dimension). It is the smallest possible vector space. We hesitate to call it $\mathbf{R}^0$, which means no components—you might think there was no vector. The vector space **Z** contains exactly *one vector* (zero). No space can do without that zero vector. Each space has its own zero vector—the zero matrix, the zero function, the vector $(0, 0, 0)$ in $\mathbf{R}^3$.

Subspaces

At different times, we will ask you to think of matrices and functions as vectors. But at all times, the vectors that we need most are ordinary column vectors. They are vectors with n components—but *maybe not all* of the vectors with n components. There are important vector spaces ***inside*** $\mathbf{R}^n$.

Start with the usual three-dimensional space $\mathbf{R}^3$. Choose a plane through the origin $(0, 0, 0)$. ***That plane is a vector space in its own right.*** If we add two vectors

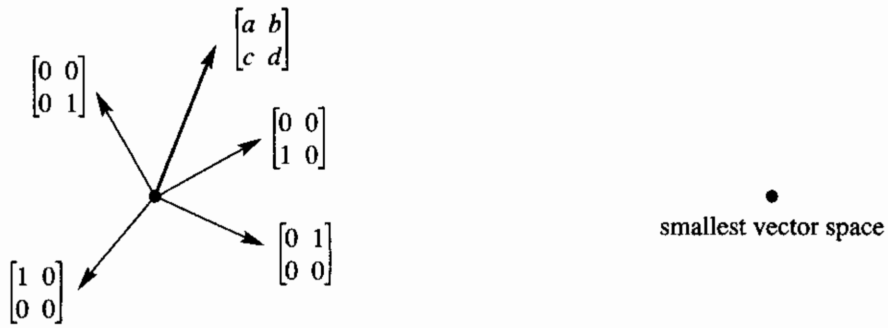


Figure 3.1 "Four-dimensional" matrix space **M**. The "zero-dimensional" space **Z**.

in the plane, their sum is in the plane. If we multiply an in-plane vector by 2 or -5, it is still in the plane. The plane is not $\mathbf{R}^2$ (even if it looks like $\mathbf{R}^2$). The vectors have three components and they belong to $\mathbf{R}^3$. The plane is a vector space inside $\mathbf{R}^3$.

This illustrates one of the most fundamental ideas in linear algebra. The plane is a ***subspace*** of the full vector space $\mathbf{R}^3$.

DEFINITION A ***subspace*** of a vector space is a set of vectors (including $\mathbf{0}$) that satisfies two requirements: ***If $\boldsymbol{v}$ and $\boldsymbol{w}$ are vectors in the subspace and c is any scalar, then*** (**i**) $\boldsymbol{v}+\boldsymbol{w}$ is in the subspace and (**ii**) $c\boldsymbol{v}$ is in the subspace.

In other words, the set of vectors is "closed" under addition $\boldsymbol{v}+\boldsymbol{w}$ and multiplication $c\boldsymbol{v}$ (and $c\boldsymbol{w}$). Those operations leave us in the subspace. We can also subtract, because $-\boldsymbol{w}$ is in the subspace and its sum with $\boldsymbol{v}$ is $\boldsymbol{v}-\boldsymbol{w}$. In short, ***all linear combinations stay in the subspace.***

All these operations follow the rules of the host space, so the eight required conditions are automatic. We just have to check the requirements (**i**) and (**ii**) for a subspace.

First fact: ***Every subspace contains the zero vector.*** The plane in $\mathbf{R}^3$ has to go through $(0,0,0)$. We mention this separately, for extra emphasis, but it follows directly from rule (**ii**). Choose $c=0$, and the rule requires $0\boldsymbol{v}$ to be in the subspace.

Planes that don't contain the origin fail those tests. When $\boldsymbol{v}$ is on such a plane, $-\boldsymbol{v}$ and $0\boldsymbol{v}$ are *not* on the plane. A plane that misses the origin is not a subspace.

Lines through the origin are also subspaces. When we multiply by 5, or add two vectors on the line, we stay on the line. But the line must go through $(0,0,0)$.

Another subspace is all of $\mathbf{R}^3$. The whole space is a subspace (*of itself*). Here is a list of all the possible subspaces of $\mathbf{R}^3$:

(**L**) Any line through $(0,0,0)$	($\mathbf{R}^3$) The whole space
(**P**) Any plane through $(0,0,0)$	(**Z**) The single vector $(0,0,0)$

If we try to keep only *part* of a plane or line, the requirements for a subspace don't hold. Look at these examples in $\mathbf{R}^2$.

Example 1 Keep only the vectors (x, y) whose components are positive or zero (this is a quarter-plane). The vector $(2, 3)$ is included but $(-2, -3)$ is not. So rule (ii) is violated when we try to multiply by $c = -1$. ***The quarter-plane is not a subspace.***

Example 2 Include also the vectors whose components are both negative. Now we have two quarter-planes. Requirement (ii) is satisfied; we can multiply by any c. But rule (i) now fails. The sum of $\boldsymbol{v} = (2, 3)$ and $\boldsymbol{w} = (-3, -2)$ is $(-1, 1)$, which is outside the quarter-planes. ***Two quarter-planes don't make a subspace.***

Rules (i) and (ii) involve vector addition $\boldsymbol{v} + \boldsymbol{w}$ and multiplication by scalars like c and d. The rules can be combined into a single requirement—*the rule for subspaces*:

A subspace containing $\boldsymbol{v}$ and $\boldsymbol{w}$ must contain all linear combinations $c\boldsymbol{v} + d\boldsymbol{w}$.

Example 3 Inside the vector space $\mathbf{M}$ of all 2 by 2 matrices, here are two subspaces:

$$(\mathbf{U}) \text{ All upper triangular matrices } \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \quad (\mathbf{D}) \text{ All diagonal matrices } \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}.$$

Add any two matrices in $\mathbf{U}$, and the sum is in $\mathbf{U}$. Add diagonal matrices, and the sum is diagonal. In this case $\mathbf{D}$ is also a subspace of $\mathbf{U}$! Of course the zero matrix is in these subspaces, when a, b, and d all equal zero.

To find a smaller subspace of diagonal matrices, we could require $a = d$. The matrices are multiples of the identity matrix I. The sum $2I + 3I$ is in this subspace, and so is 3 times $4I$. It is a "line of matrices" inside $\mathbf{M}$ and $\mathbf{U}$ and $\mathbf{D}$.

Is the matrix I a subspace by itself? Certainly not. Only the zero matrix is. Your mind will invent more subspaces of 2 by 2 matrices—write them down for Problem 5.

The Column Space of A

The most important subspaces are tied directly to a matrix A. We are trying to solve $A\boldsymbol{x} = \boldsymbol{b}$. If A is not invertible, the system is solvable for some $\boldsymbol{b}$ and not solvable for other $\boldsymbol{b}$. We want to describe the good right sides $\boldsymbol{b}$—the vectors that *can* be written as A times some vector $\boldsymbol{x}$.

Remember that $A\boldsymbol{x}$ is a combination of the columns of A. To get every possible $\boldsymbol{b}$, we use every possible $\boldsymbol{x}$. So start with the columns of A, and ***take all their linear combinations. This produces the column space of A.*** It is a vector space made up of column vectors—not just the n columns of A, but all their combinations $A\boldsymbol{x}$.

DEFINITION The ***column space*** consists of ***all linear combinations of the columns.*** The combinations are all possible vectors $A\boldsymbol{x}$. They fill the column space $\boldsymbol{C}(A)$.

This column space is crucial to the whole book, and here is why. ***To solve $A\boldsymbol{x} = \boldsymbol{b}$ is to express $\boldsymbol{b}$ as a combination of the columns.*** The right side $\boldsymbol{b}$ has to be *in the column space* produced by A on the left side, or no solution!

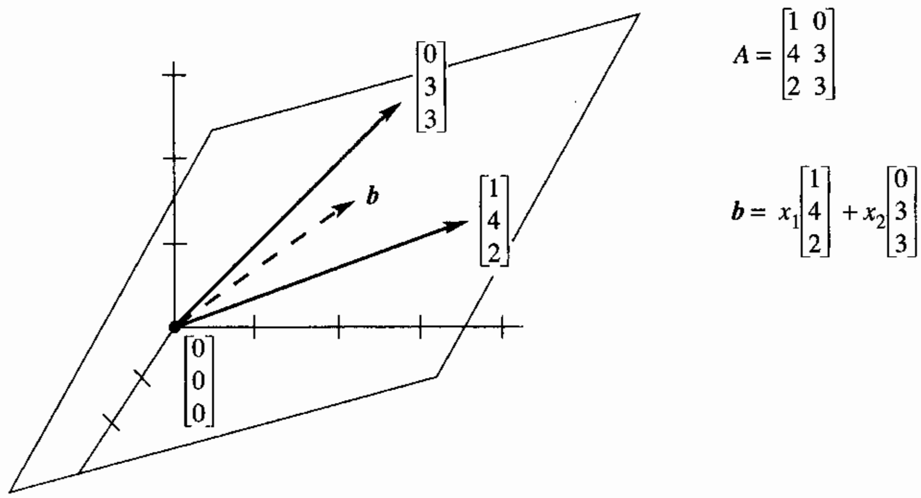


Figure 3.2 The column space $C(A)$ is a plane containing the two columns. $A\boldsymbol{x} = \boldsymbol{b}$ is solvable when $\boldsymbol{b}$ is on that plane. Then $\boldsymbol{b}$ is a combination of the columns.

3A The system $A\boldsymbol{x} = \boldsymbol{b}$ is solvable if and only if $\boldsymbol{b}$ is in the column space of A.

When $\boldsymbol{b}$ is in the column space, it is a combination of the columns. The coefficients in that combination give us a solution $\boldsymbol{x}$ to the system $A\boldsymbol{x} = \boldsymbol{b}$.

Suppose A is an m by n matrix. Its columns have m components (not n). So the columns belong to $\mathbf{R}^m$. ***The column space of A is a subspace of*** $\mathbf{R}^m$ ***(not*** $\mathbf{R}^n$***).*** The set of all column combinations $A\boldsymbol{x}$ satisfies rules (i) and (ii) for a subspace: When we add linear combinations or multiply by scalars, we still produce combinations of the columns. The word "subspace" is justified *by taking all linear combinations.*

Here is a 3 by 2 matrix A, whose column space is a subspace of $\mathbf{R}^3$. It is a plane.

Example 4

$$A\boldsymbol{x} \quad \text{is} \quad \begin{bmatrix}1 & 0\\4 & 3\\2 & 3\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} \quad \text{which is} \quad x_1\begin{bmatrix}1\\4\\2\end{bmatrix} + x_2\begin{bmatrix}0\\3\\3\end{bmatrix}.$$

The column space consists of all combinations of the two columns—any x_1 times the first column plus any x_2 times the second column. *Those combinations fill up a plane in* $\mathbf{R}^3$ (Figure 3.2). If the right side $\boldsymbol{b}$ lies on that plane, then it is one of the combinations and (x_1, x_2) is a solution to $A\boldsymbol{x} = \boldsymbol{b}$. The plane has zero thickness, so it is more likely that $\boldsymbol{b}$ is not in the column space. Then there is no solution to our 3 equations in 2 unknowns.

Of course $(0, 0, 0)$ is in the column space. The plane passes through the origin. There is certainly a solution to $A\boldsymbol{x} = \boldsymbol{0}$. That solution, always available, is $\boldsymbol{x} =$ _____.

To repeat, the attainable right sides $\boldsymbol{b}$ are exactly the vectors in the column space. One possibility is the first column itself—take $x_1 = 1$ and $x_2 = 0$. Another combination is the second column—take $x_1 = 0$ and $x_2 = 1$. The new level of understanding is to see *all* combinations—the whole subspace is generated by those two columns.

Notation The column space of A is denoted by $\boldsymbol{C}(A)$. Start with the columns and take all their linear combinations. We might get the whole $\mathbf{R}^m$ or only a subspace.

Example 5 Describe the column spaces (they are subspaces of $\mathbf{R}^2$) for

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}.$$

Solution The column space of I is the *whole space* $\mathbf{R}^2$. Every vector is a combination of the columns of I. In vector space language, $\boldsymbol{C}(I)$ is $\mathbf{R}^2$.

The column space of A is only a line. The second column $(2, 4)$ is a multiple of the first column $(1, 2)$. Those vectors are different, but our eye is on vector ***spaces***. The column space contains $(1, 2)$ and $(2, 4)$ and all other vectors $(c, 2c)$ along that line. The equation $A\boldsymbol{x} = \boldsymbol{b}$ is only solvable when $\boldsymbol{b}$ is on the line.

The third matrix (with three columns) places no restriction on $\boldsymbol{b}$. The column space $\boldsymbol{C}(B)$ is all of $\mathbf{R}^2$. Every $\boldsymbol{b}$ is attainable. The vector $\boldsymbol{b} = (5, 4)$ is column 2 plus column 3, so $\boldsymbol{x}$ can be $(0, 1, 1)$. The same vector $(5, 4)$ is also 2(column 1) + column 3, so another possible $\boldsymbol{x}$ is $(2, 0, 1)$. This matrix has the same column space as I—any $\boldsymbol{b}$ is allowed. But now $\boldsymbol{x}$ has extra components and there are more solutions.

The next section creates another vector space, to describe all the solutions of $A\boldsymbol{x} = \mathbf{0}$. This section created the column space, to describe all the attainable right sides $\boldsymbol{b}$.

■ REVIEW OF THE KEY IDEAS ■

1. $\mathbf{R}^n$ contains all column vectors with n real components.
2. $\mathbf{M}$ (2 by 2 matrices) and $\mathbf{F}$ (functions) and $\mathbf{Z}$ (zero vector alone) are vector spaces.
3. A subspace containing $\boldsymbol{v}$ and $\boldsymbol{w}$ must contain all their combinations $c\boldsymbol{v} + d\boldsymbol{w}$.
4. The combinations of the columns of A form the ***column space*** $\boldsymbol{C}(A)$.
5. $A\boldsymbol{x} = \boldsymbol{b}$ has a solution exactly when $\boldsymbol{b}$ is in the column space of A.

■ WORKED EXAMPLES ■

3.1 A We are given three different vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$. Construct a matrix so that the equations $A\boldsymbol{x} = \boldsymbol{b}_1$ and $A\boldsymbol{x} = \boldsymbol{b}_2$ are solvable, but $A\boldsymbol{x} = \boldsymbol{b}_3$ is not solvable. How can you decide if this is possible? How could you construct A?

Solution We want to have $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ in the column space of A. Then $A\boldsymbol{x} = \boldsymbol{b}_1$ and $A\boldsymbol{x} = \boldsymbol{b}_2$ will be solvable. *The quickest way is to make $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ the two columns of* A. Then the solutions are $\boldsymbol{x} = (1, 0)$ and $\boldsymbol{x} = (0, 1)$.

Also, we don't want $A\boldsymbol{x} = \boldsymbol{b}_3$ to be solvable. So don't make the column space any larger! Keeping only the columns of $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, the question is:

$$\text{Is } A\boldsymbol{x} = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \boldsymbol{b}_3 \text{ solvable?} \qquad \text{Is } \boldsymbol{b}_3 \text{ a combination of } \boldsymbol{b}_1 \text{ and } \boldsymbol{b}_2\text{?}$$

If the answer is *no*, we have the desired matrix A. If the answer is *yes*, then it is *not possible* to construct A. When the column space contains $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, it will have to contain all their linear combinations. So $\boldsymbol{b}_3$ would necessarily be in that column space and $A\boldsymbol{x} = \boldsymbol{b}_3$ would necessarily be solvable.

3.1 B Describe a subspace **S** of each vector space **V**, and then a subspace **SS** of **S**.

$\mathbf{V}_1$ = all combinations of $(1, 1, 0, 0)$ and $(1, 1, 1, 0)$ and $(1, 1, 1, 1)$
$\mathbf{V}_2$ = all vectors perpendicular to $\boldsymbol{u} = (1, 2, 2, 1)$
$\mathbf{V}_3$ = all symmetric 2 by 2 matrices
$\mathbf{V}_4$ = all solutions to the equation $d^4y/dx^4 = 0$.

Describe each **V** two ways: *All combinations of ...., all solutions of the equations ....*

Solution A subspace **S** of $\mathbf{V}_1$ comes from all combinations of the first two vectors $(1, 1, 0, 0)$ and $(1, 1, 1, 0)$. A subspace **SS** of **S** comes from all multiples $(c, c, 0, 0)$ of the first vector.

A subspace **S** of $\mathbf{V}_2$ comes from all combinations of two vectors $(1, 0, 0, -1)$ and $(0, 1, -1, 0)$ that are perpendicular to $\boldsymbol{u}$. The vector $\boldsymbol{x} = (1, 1, -1, -1)$ is in **S** and all its multiples $c\boldsymbol{x}$ give a subspace **SS**.

The diagonal matrices are a subspace **S** of the symmetric matrices. The multiples cI are a subspace **SS** of the diagonal matrices.

$\mathbf{V}_4$ contains all cubic polynomials $y = a + bx + cx^2 + dx^3$. The quadratic polynomials give a subspace **S**. The linear polynomials are one choice of **SS**. The constants could be **SSS**.

In all four parts we could have chosen **S** = **V** itself, and **SS** = the zero subspace **Z**.

Each **V** can be described as all combinations of .... and as all solutions of ....:

$\mathbf{V}_1$ = all combinations of the 3 vectors = all solutions of $v_1 - v_2 = 0$
$\mathbf{V}_2$ = all combinations of $(1, 0, 0, -1), (0, 1, -1, 0), (2, -1, 0, 0)$
= all solutions of $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = 0$
$\mathbf{V}_3$ = all combinations of $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ = all solutions $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ of $b = c$
$\mathbf{V}_4$ = all combinations of $1, x, x^2, x^3$ = all solutions to $d^4y/dx^4 = 0$.

Problem Set 3.1

The first problems 1–8 are about vector spaces in general. The vectors in those spaces are not necessarily column vectors. In the definition of a *vector space*, vector addition $x+y$ and scalar multiplication cx must obey the following eight rules:

(1) $x + y = y + x$

(2) $x + (y + z) = (x + y) + z$

(3) There is a unique "zero vector" such that $x + 0 = x$ for all x

(4) For each x there is a unique vector $-x$ such that $x + (-x) = 0$

(5) 1 times x equals x

(6) $(c_1c_2)x = c_1(c_2x)$

(7) $c(x + y) = cx + cy$

(8) $(c_1 + c_2)x = c_1x + c_2x$.

1 Suppose $(x_1, x_2) + (y_1, y_2)$ is defined to be $(x_1 + y_2, x_2 + y_1)$. With the usual multiplication $cx = (cx_1, cx_2)$, which of the eight conditions are not satisfied?

2 Suppose the multiplication cx is defined to produce $(cx_1, 0)$ instead of (cx_1, cx_2). With the usual addition in $\mathbf{R}^2$, are the eight conditions satisfied?

3 (a) Which rules are broken if we keep only the positive numbers $x > 0$ in $\mathbf{R}^1$? Every c must be allowed. The half-line is not a subspace.

(b) The positive numbers with $x + y$ and cx redefined to equal the usual xy and x^c *do* satisfy the eight rules. Test rule 7 when $c = 3, x = 2, y = 1$. (Then $x + y = 2$ and $cx = 8$.) Which number acts as the "zero vector"?

4 The matrix $A = \begin{bmatrix} 2 & -2 \\ 2 & -2 \end{bmatrix}$ is a "vector" in the space $\mathbf{M}$ of all 2 by 2 matrices. Write down the zero vector in this space, the vector $\frac{1}{2}A$, and the vector $-A$. What matrices are in the smallest subspace containing A?

5 (a) Describe a subspace of $\mathbf{M}$ that contains $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ but not $B = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$.

(b) If a subspace of $\mathbf{M}$ contains A and B, must it contain I?

(c) Describe a subspace of $\mathbf{M}$ that contains no nonzero diagonal matrices.

6 The functions $f(x) = x^2$ and $g(x) = 5x$ are "vectors" in $\mathbf{F}$. This is the vector space of all real functions. (The functions are defined for $-\infty < x < \infty$.) The combination $3f(x) - 4g(x)$ is the function $h(x) =$ ______.

7 Which rule is broken if multiplying $f(x)$ by c gives the function $f(cx)$? Keep the usual addition $f(x) + g(x)$.

8 If the sum of the "vectors" $f(x)$ and $g(x)$ is defined to be the function $f(g(x))$, then the "zero vector" is $g(x) = x$. Keep the usual scalar multiplication $cf(x)$ and find two rules that are broken.

Questions 9–18 are about the "subspace requirements": $\boldsymbol{x} + \boldsymbol{y}$ and $c\boldsymbol{x}$ (and then all linear combinations $c\boldsymbol{x} + d\boldsymbol{y}$) must stay in the subspace.

9 One requirement can be met while the other fails. Show this by finding

(a) A set of vectors in $\mathbf{R}^2$ for which $\boldsymbol{x} + \boldsymbol{y}$ stays in the set but $\frac{1}{2}\boldsymbol{x}$ may be outside.

(b) A set of vectors in $\mathbf{R}^2$ (other than two quarter-planes) for which every $c\boldsymbol{x}$ stays in the set but $\boldsymbol{x} + \boldsymbol{y}$ may be outside.

10 Which of the following subsets of $\mathbf{R}^3$ are actually subspaces?

(a) The plane of vectors (b_1, b_2, b_3) with $b_1 = b_2$.

(b) The plane of vectors with $b_1 = 1$.

(c) The vectors with $b_1 b_2 b_3 = 0$.

(d) All linear combinations of $\boldsymbol{v} = (1, 4, 0)$ and $\boldsymbol{w} = (2, 2, 2)$.

(e) All vectors that satisfy $b_1 + b_2 + b_3 = 0$.

(f) All vectors with $b_1 \le b_2 \le b_3$.

11 Describe the smallest subspace of the matrix space $\mathbf{M}$ that contains

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

12 Let P be the plane in $\mathbf{R}^3$ with equation $x + y - 2z = 4$. The origin $(0, 0, 0)$ is not in P! Find two vectors in P and check that their sum is not in P.

13 Let $\mathbf{P}_0$ be the plane through $(0, 0, 0)$ parallel to the previous plane P. What is the equation for $\mathbf{P}_0$? Find two vectors in $\mathbf{P}_0$ and check that their sum is in $\mathbf{P}_0$.

14 The subspaces of $\mathbf{R}^3$ are planes, lines, $\mathbf{R}^3$ itself, or $\mathbf{Z}$ containing only $(0, 0, 0)$.

(a) Describe the three types of subspaces of $\mathbf{R}^2$.

(b) Describe the five types of subspaces of $\mathbf{R}^4$.

15 (a) The intersection of two planes through $(0, 0, 0)$ is probably a _____ but it could be a _____. It can't be $\mathbf{Z}$!

(b) The intersection of a plane through $(0,0,0)$ with a line through $(0,0,0)$ is probably a ____ but it could be a ____.

(c) If $\mathbf{S}$ and $\mathbf{T}$ are subspaces of $\mathbf{R}^5$, prove that their intersection $\mathbf{S} \cap \mathbf{T}$ (vectors in both subspaces) is a subspace of $\mathbf{R}^5$. *Check the requirements on $x+y$ and cx.*

16 Suppose $\mathbf{P}$ is a plane through $(0,0,0)$ and $\mathbf{L}$ is a line through $(0,0,0)$. The smallest vector space containing both $\mathbf{P}$ and $\mathbf{L}$ is either ____ or ____.

17 (a) Show that the set of *invertible* matrices in $\mathbf{M}$ is not a subspace.

(b) Show that the set of *singular* matrices in $\mathbf{M}$ is not a subspace.

18 True or false (check addition in each case by an example):

(a) The symmetric matrices in $\mathbf{M}$ (with $A^{\mathrm{T}} = A$) form a subspace.

(b) The skew-symmetric matrices in $\mathbf{M}$ (with $A^{\mathrm{T}} = -A$) form a subspace.

(c) The unsymmetric matrices in $\mathbf{M}$ (with $A^{\mathrm{T}} \neq A$) form a subspace.

Questions 19–27 are about column spaces $C(A)$ and the equation $Ax = b$.

19 Describe the column spaces (lines or planes) of these particular matrices:

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

20 For which right sides (find a condition on b_1, b_2, b_3) are these systems solvable?

$$\text{(a)} \quad \begin{bmatrix} 1 & 4 & 2 \\ 2 & 8 & 4 \\ -1 & -4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \qquad \text{(b)} \quad \begin{bmatrix} 1 & 4 \\ 2 & 9 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

21 Adding row 1 of A to row 2 produces B. Adding column 1 to column 2 produces C. A combination of the columns of ____ is also a combination of the columns of A. Which two matrices have the same column ____?

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}.$$

22 For which vectors (b_1, b_2, b_3) do these systems have a solution?

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\text{and} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

23 (Recommended) If we add an extra column $\boldsymbol{b}$ to a matrix A, then the column space gets larger unless _____. Give an example where the column space gets larger and an example where it doesn't. Why is $A\boldsymbol{x} = \boldsymbol{b}$ solvable exactly when the column space *doesn't* get larger—it is the same for A and $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$?

24 The columns of AB are combinations of the columns of A. This means: *The column space of AB is contained in* (possibly equal to) *the column space of A.* Give an example where the column spaces of A and AB are not equal.

25 Suppose $A\boldsymbol{x} = \boldsymbol{b}$ and $A\boldsymbol{y} = \boldsymbol{b}^*$ are both solvable. Then $A\boldsymbol{z} = \boldsymbol{b} + \boldsymbol{b}^*$ is solvable. What is $\boldsymbol{z}$? This translates into: If $\boldsymbol{b}$ and $\boldsymbol{b}^*$ are in the column space $\boldsymbol{C}(A)$, then $\boldsymbol{b} + \boldsymbol{b}^*$ is in $\boldsymbol{C}(A)$.

26 If A is any 5 by 5 invertible matrix, then its column space is _____. Why?

27 True or false (with a counterexample if false):

(a) The vectors $\boldsymbol{b}$ that are not in the column space $\boldsymbol{C}(A)$ form a subspace.

(b) If $\boldsymbol{C}(A)$ contains only the zero vector, then A is the zero matrix.

(c) The column space of $2A$ equals the column space of A.

(d) The column space of $A - I$ equals the column space of A.

28 Construct a 3 by 3 matrix whose column space contains $(1, 1, 0)$ and $(1, 0, 1)$ but not $(1, 1, 1)$. Construct a 3 by 3 matrix whose column space is only a line.

29 If the 9 by 12 system $A\boldsymbol{x} = \boldsymbol{b}$ is solvable for every $\boldsymbol{b}$, then $\boldsymbol{C}(A) =$ _____.

THE NULLSPACE OF A: SOLVING $AX = 0$ ■ 3.2

This section is about the space of solutions to $A\boldsymbol{x} = \mathbf{0}$. The matrix A can be square or rectangular. *One immediate solution is* $\boldsymbol{x} = \mathbf{0}$. For invertible matrices this is the only solution. For other matrices, not invertible, there are nonzero solutions to $A\boldsymbol{x} = \mathbf{0}$. *Each solution* $\boldsymbol{x}$ *belongs to the nullspace of* A.

Elimination will find all solutions and identify this very important subspace.

DEFINITION ***The nullspace of*** A ***consists of all solutions to*** $A\boldsymbol{x} = \mathbf{0}$**.** These solution vectors $\boldsymbol{x}$ are in $\mathbf{R}^n$. The nullspace containing all solutions is denoted by $\boldsymbol{N}(A)$.

Check that the solution vectors form a subspace. Suppose $\boldsymbol{x}$ and $\boldsymbol{y}$ are in the nullspace (this means $A\boldsymbol{x} = \mathbf{0}$ and $A\boldsymbol{y} = \mathbf{0}$). The rules of matrix multiplication give $A(\boldsymbol{x} + \boldsymbol{y}) = \mathbf{0} + \mathbf{0}$. The rules also give $A(c\boldsymbol{x}) = c\mathbf{0}$. The right sides are still zero. Therefore $\boldsymbol{x} + \boldsymbol{y}$ and $c\boldsymbol{x}$ are also in the nullspace $\boldsymbol{N}(A)$. Since we can add and multiply without leaving the nullspace, it is a subspace.

To repeat: The solution vectors $\boldsymbol{x}$ have n components. They are vectors in $\mathbf{R}^n$, so *the nullspace is a subspace of* $\mathbf{R}^n$. The column space $\boldsymbol{C}(A)$ is a subspace of $\mathbf{R}^m$.

If the right side $\boldsymbol{b}$ is not zero, the solutions of $A\boldsymbol{x} = \boldsymbol{b}$ do *not* form a subspace. The vector $\boldsymbol{x} = \mathbf{0}$ is only a solution if $\boldsymbol{b} = \mathbf{0}$. When the set of solutions does not include $\boldsymbol{x} = \mathbf{0}$, it cannot be a subspace. Section 3.4 will show how the solutions to $A\boldsymbol{x} = \boldsymbol{b}$ (if there are any solutions) are shifted away from the origin by one particular solution.

Example 1 The equation $x + 2y + 3z = 0$ comes from the 1 by 3 matrix $A = [\,1\ \ 2\ \ 3\,]$. This equation produces a plane through the origin. The plane is a subspace of $\mathbf{R}^3$. *It is the nullspace of* A.

The solutions to $x + 2y + 3z = 6$ also form a plane, but not a subspace.

Example 2 Describe the nullspace of $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$.

Solution Apply elimination to the linear equations $A\boldsymbol{x} = \mathbf{0}$:

$$\begin{bmatrix} x_1 + 2x_2 = 0 \\ 3x_1 + 6x_2 = 0 \end{bmatrix} \quad \to \quad \begin{bmatrix} x_1 + 2x_2 = 0 \\ 0 = 0 \end{bmatrix}$$

There is really only one equation. The second equation is the first equation multiplied by 3. In the row picture, the line $x_1 + 2x_2 = 0$ is the same as the line $3x_1 + 6x_2 = 0$. That line is the nullspace $\boldsymbol{N}(A)$.

To describe this line of solutions, here is an efficient way. Choose one point on the line (one "***special solution***"). Then all points on the line are multiples of this one. We choose the second component to be $x_2 = 1$ (a special choice). From the equation $x_1 + 2x_2 = 0$, the first component must be $x_1 = -2$. The special solution is $(-2, 1)$:

$$\text{The nullspace } \boldsymbol{N}(A) \text{ contains all multiples of } \boldsymbol{s} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

This is the best way to describe the nullspace, by computing special solutions to $Ax = \mathbf{0}$. ***The nullspace consists of all combinations of those special solutions.*** This example has one special solution and the nullspace is a line.

The plane $x + 2y + 3z = 0$ in Example 1 had *two* special solutions:

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \text{ has the special solutions } s_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \text{ and } s_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

Those vectors s_1 and s_2 lie on the plane $x + 2y + 3z = 0$, which is the nullspace of $A = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$. All vectors on the plane are combinations of s_1 and s_2.

Notice what is special about s_1 and s_2. They have ones and zeros in the last two components. *Those components are "free" and we choose them specially.* Then the first components -2 and -3 are determined by the equation $Ax = \mathbf{0}$.

The first column of $A = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ contains the *pivot*, so the first component of x is *not free*. The free components correspond to columns without pivots. This description of special solutions will be completed after one more example.

The special choice (one or zero) is only for the free variables.

Example 3 Describe the nullspaces of these three matrices A, B, C:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix} \quad B = \begin{bmatrix} A \\ 2A \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 4 \\ 6 & 16 \end{bmatrix} \quad C = \begin{bmatrix} A & 2A \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 3 & 8 & 6 & 16 \end{bmatrix}.$$

Solution The equation $Ax = \mathbf{0}$ has only the zero solution $x = \mathbf{0}$. *The nullspace is* $\mathbf{Z}$. It contains only the single point $x = \mathbf{0}$ in $\mathbf{R}^2$. This comes from elimination:

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ yields } \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} x_1 = 0 \\ x_2 = 0 \end{bmatrix}.$$

A is invertible. There are no special solutions. All columns have pivots.

The rectangular matrix B has the same nullspace $\mathbf{Z}$. The first two equations in $Bx = \mathbf{0}$ again require $x = \mathbf{0}$. The last two equations would also force $x = \mathbf{0}$. When we add extra equations, the nullspace certainly cannot become larger. The extra rows impose more conditions on the vectors x in the nullspace.

The rectangular matrix C is different. It has extra columns instead of extra rows. The solution vector x has *four* components. Elimination will produce pivots in the first two columns of C, but the last two columns are "free". They don't have pivots:

$$C = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 3 & 8 & 6 & 16 \end{bmatrix} \text{ becomes } U = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 2 & 0 & 4 \end{bmatrix}$$
$$\begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ \multicolumn{2}{c}{\text{pivot columns}} & \multicolumn{2}{c}{\text{free columns}} \end{array}$$

For the free variables x_3 and x_4, we make special choices of ones and zeros. First $x_3 = 1$, $x_4 = 0$ and second $x_3 = 0$, $x_4 = 1$. The pivot variables x_1 and x_2 are determined by the equation $Ux = \mathbf{0}$. We get two special solutions in the nullspace of C (and also the nullspace of U). The special solutions are:

$$s_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } s_2 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix} \begin{matrix} \leftarrow \text{ pivot} \\ \leftarrow \quad \text{variables} \\ \leftarrow \text{ free} \\ \leftarrow \quad \text{variables} \end{matrix}$$

One more comment to anticipate what is coming soon. Elimination will not stop at the upper triangular U! We can continue to make this matrix simpler, in two ways:

1. ***Produce zeros above the pivots***, by eliminating upward.

2. ***Produce ones in the pivots***, by dividing the whole row by its pivot.

Those steps don't change the zero vector on the right side of the equation. The nullspace stays the same. This nullspace becomes easiest to see when we reach the ***reduced row echelon form*** R. It has I in the pivot columns:

$$U = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 2 & 0 & 4 \end{bmatrix} \quad \text{becomes} \quad R = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}.$$
$$\uparrow \; \uparrow$$
pivot columns contain I

I subtracted row 2 of U from row 1, and then multiplied row 2 by $\frac{1}{2}$. The original two equations have simplified to $x_1 + 2x_3 = 0$ and $x_2 + 2x_4 = 0$.

The first special solution is still $s_1 = (-2, 0, 1, 0)$, and s_2 is unchanged. Special solutions are much easier to find from the reduced system $Rx = \mathbf{0}$.

Before moving to m by n matrices A and their nullspaces $N(A)$ and special solutions, allow me to repeat one comment. For many matrices, the only solution to $Ax = \mathbf{0}$ is $x = \mathbf{0}$. Their nullspaces $N(A) = \mathbf{Z}$ contain only that one vector. The only combination of the columns that produces $b = \mathbf{0}$ is then the "zero combination" or "trivial combination". The solution is trivial (just $x = \mathbf{0}$) but the idea is not trivial.

This case of a zero nullspace $\mathbf{Z}$ is of the greatest importance. It says that the columns of A are independent. No combination of columns gives the zero vector (except the zero combination). All columns have pivots and no columns are free. You will see this idea of independence again ...

Solving $Ax = \mathbf{0}$ by Elimination

This is important. ***A is rectangular and we still use elimination***. We solve m equations in n unknowns when $b = 0$. After A is simplified by row operations, we read off the solution (or solutions). Remember the two stages in solving $Ax = \mathbf{0}$:

1. Forward elimination from A to a triangular U (or its reduced form R).

2. Back substitution in $Ux = \mathbf{0}$ or $Rx = \mathbf{0}$ to find x.

You will notice a difference in back substitution, when A and U have fewer than n pivots. *We are allowing all matrices in this chapter*, not just the nice ones (which are square matrices with inverses).

Pivots are still nonzero. The columns below the pivots are still zero. But it might happen that a column has no pivot. In that case, don't stop the calculation. ***Go on to the next column***. The first example is a 3 by 4 matrix with two pivots:

$$A = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 8 & 10 \\ 3 & 3 & 10 & 13 \end{bmatrix}.$$

Certainly $a_{11} = 1$ is the first pivot. Clear out the 2 and 3 below that pivot:

$$A \to \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 4 & 4 \end{bmatrix} \qquad \begin{matrix} \\ \text{(subtract } 2 \times \text{ row 1)} \\ \text{(subtract } 3 \times \text{ row 1)} \end{matrix}$$

The second column has a zero in the pivot position. We look below the zero for a nonzero entry, ready to do a row exchange. *The entry below that position is also zero.* Elimination can do nothing with the second column. This signals trouble, which we expect anyway for a rectangular matrix. There is no reason to quit, and we go on to the third column.

The second pivot is 4 (but it is in the third column). Subtracting row 2 from row 3 clears out that column below the pivot. We arrive at

$$\textbf{Triangular } U: \quad U = \begin{bmatrix} \mathbf{1} & 1 & 2 & 3 \\ 0 & 0 & \mathbf{4} & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{matrix} \textit{(only two pivots)} \\ \textit{(the last equation} \\ \textit{became } 0 = 0) \end{matrix}$$

The fourth column also has a zero in the pivot position—but nothing can be done. There is no row below it to exchange, and forward elimination is complete. The matrix has three rows, four columns, and *only two pivots.* The original $Ax = \mathbf{0}$ seemed to involve three different equations, but the third equation is the sum of the first two. It is automatically satisfied $(0 = 0)$ when the first two equations are satisfied. Elimination reveals the inner truth about a system of equations. Soon we push on from U to R.

Now comes back substitution, to find all solutions to $Ux = \mathbf{0}$. With four unknowns and only two pivots, there are many solutions. The question is how to write them all down. A good method is to separate the ***pivot variables*** from the ***free variables***.

P The ***pivot*** variables are x_1 and x_3, since columns 1 and 3 contain pivots.
F The ***free*** variables are x_2 and x_4, because columns 2 and 4 have no pivots.

The free variables x_2 and x_4 can be given any values whatsoever. Then back substitution finds the pivot variables x_1 and x_3. (In Chapter 2 no variables were free. When A is invertible, all variables are pivot variables.) The simplest choices for the free variables are ones and zeros. Those choices give the *special solutions.*

Special Solutions to $x_1 + x_2 + 2x_3 + 3x_4 = 0$ and $4x_3 + 4x_4 = 0$

- Set $x_2 = 1$ and $x_4 = 0$. By back substitution $x_3 = 0$. Then $x_1 = -1$.
- Set $x_2 = 0$ and $x_4 = 1$. By back substitution $x_3 = -1$. Then $x_1 = -1$.

These special solutions solve $U\boldsymbol{x} = \mathbf{0}$ and therefore $A\boldsymbol{x} = \mathbf{0}$. They are in the nullspace. The good thing is that *every solution is a combination of the special solutions.*

$$\textbf{\textit{Complete Solution}} \quad \boldsymbol{x} = \underset{\textbf{special}}{x_2\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}} + \underset{\textbf{special}}{x_4\begin{bmatrix}-1\\0\\-1\\1\end{bmatrix}} = \underset{\textbf{complete}}{\begin{bmatrix}-x_2-x_4\\x_2\\-x_4\\x_4\end{bmatrix}}. \tag{1}$$

Please look again at that answer. It is the main goal of this section. The vector $\boldsymbol{s}_1 = (-1, 1, 0, 0)$ is the special solution when $x_2 = 1$ and $x_4 = 0$. The second special solution has $x_2 = 0$ and $x_4 = 1$. ***All solutions are linear combinations of*** $\boldsymbol{s}_1$ ***and*** $\boldsymbol{s}_2$. The special solutions are in the nullspace $\boldsymbol{N}(A)$, and their combinations fill out the whole nullspace.

The MATLAB code **nullbasis** computes these special solutions. They go into the columns of a ***nullspace matrix*** N. The complete solution to $A\boldsymbol{x} = \mathbf{0}$ is a combination of those columns. Once we have the special solutions, we have the whole nullspace.

There is a special solution for each free variable. If no variables are free—this means there are n pivots—then the only solution to $U\boldsymbol{x} = \mathbf{0}$ and $A\boldsymbol{x} = \mathbf{0}$ is the trivial solution $\boldsymbol{x} = \mathbf{0}$. All variables are pivot variables. In that case the nullspaces of A and U contain only the zero vector. With no free variables, and pivots in every column, the output from **nullbasis** is an empty matrix.

Example 4 Find the nullspace of $U = \begin{bmatrix}1&5&7\\0&0&9\end{bmatrix}$.

The second column of U has no pivot. So x_2 is free. The special solution has $x_2 = 1$. Back substitution into $9x_3 = 0$ gives $x_3 = 0$. Then $x_1 + 5x_2 = 0$ or $x_1 = -5$. The solutions to $U\boldsymbol{x} = \mathbf{0}$ are multiples of one special solution:

$$\boldsymbol{x} = x_2\begin{bmatrix}-5\\1\\0\end{bmatrix}$$

The nullspace of U is a line in $\mathbf{R}^3$.
It contains multiples of the special solution.
One variable is free, and $N = $ **nullbasis** (U) has one column.

In a minute we will continue elimination on U, to get *zeros above the pivots and ones in the pivots*. The 7 is eliminated and the pivot changes from 9 to 1. The final result of this elimination will be the reduced row echelon matrix R:

$$U = \begin{bmatrix} 1 & 5 & 7 \\ 0 & 0 & 9 \end{bmatrix} \text{ reduces to } R = \begin{bmatrix} 1 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This makes it even clearer that the special solution (column of N) is $\boldsymbol{s} = (-5, 1, 0)$.

Echelon Matrices

Forward elimination goes from A to U. The process starts with an m by n matrix A. It acts by row operations, including row exchanges. It goes on to the next column when no pivot is available in the current column. The m by n "staircase," U is an ***echelon matrix***.

Here is a 4 by 7 echelon matrix with the three pivots highlighted in boldface:

$$U = \begin{bmatrix} \boldsymbol{x} & x & x & x & x & x & x \\ 0 & \boldsymbol{x} & x & x & x & x & x \\ 0 & 0 & 0 & 0 & 0 & \boldsymbol{x} & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{array}{l} \text{Three pivot variables } x_1, x_2, x_6 \\ \text{Four free variables } x_3, x_4, x_5, x_7 \\ \text{Four special solutions in } N(U) \end{array}$$

Question What are the column space and the nullspace for this matrix?

Answer The columns have four components so they lie in $\mathbf{R}^4$. (Not in $\mathbf{R}^3$!) The fourth component of every column is zero. Every combination of the columns—every vector in the column space—has fourth component zero. *The column space $C(U)$ consists of all vectors of the form* $(b_1, b_2, b_3, 0)$. For those vectors we can solve $U\boldsymbol{x} = \boldsymbol{b}$ by back substitution. These vectors $\boldsymbol{b}$ are all possible combinations of the seven columns.

The nullspace $N(U)$ is a subspace of $\mathbf{R}^7$. The solutions to $U\boldsymbol{x} = \mathbf{0}$ are all the combinations of the four special solutions—*one for each free variable*:

1. Columns $3, 4, 5, 7$ have no pivots. So the free variables are x_3, x_4, x_5, x_7.

2. Set one free variable to 1 and set the other free variables to zero.

3. Solve $U\boldsymbol{x} = \mathbf{0}$ for the pivot variables x_1, x_2, x_6.

4. This gives one of the four special solutions in the nullspace matrix N.

The nonzero rows of an echelon matrix go down in a staircase pattern. The pivots are the first nonzero entries in those rows. There is a column of zeros below every pivot.

Counting the pivots leads to an extremely important theorem. Suppose A has more columns than rows. ***With $n > m$ there is at least one free variable.*** The system $A\boldsymbol{x} = \mathbf{0}$ has at least one special solution. This solution is *not zero*!

3B If $A\boldsymbol{x} = \mathbf{0}$ has more unknowns than equations ($n > m$, more columns than rows), then it has nonzero solutions. There must be free columns, without pivots.

In other words, a short wide matrix ($n > m$) always has nonzero vectors in its nullspace. There must be at least $n-m$ free variables, since the number of pivots cannot exceed m. (The matrix only has m rows, and a row never has two pivots.) Of course a row might have *no* pivot—which means an extra free variable. But here is the point: When there is a free variable, it can be set to 1. Then the equation $A\boldsymbol{x} = \mathbf{0}$ has a nonzero solution.

To repeat: There are at most m pivots. With $n > m$, the system $A\boldsymbol{x} = \mathbf{0}$ has a nonzero solution. Actually there are infinitely many solutions, since any multiple $c\boldsymbol{x}$ is also a solution. The nullspace contains at least a line of solutions. With two free variables, there are two special solutions and the nullspace is even larger.

The nullspace is a subspace. Its "dimension" is the number of free variables. This central idea—the ***dimension*** of a subspace—is defined and explained in this chapter.

The Reduced Row Echelon Matrix R

>From an echelon matrix U we can go one more step. Continue with our example

$$U = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We can divide the second row by 4. Then both pivots equal 1. *We can subtract* 2 *times this new row* $\begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}$ *from the row above.* The ***reduced row echelon matrix R has zeros above the pivots as well as below***:

$$R = \begin{bmatrix} \mathbf{1} & 1 & 0 & 1 \\ 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

R has 1's as pivots. Zeros above pivots come from ***upward elimination***.

If A is invertible, its reduced row echelon form is the identity matrix $R = I$. This is the ultimate in row reduction. Of course the nullspace is then $\mathbf{Z}$.

The zeros in R make it easy to find the special solutions (the same as before):

1. Set $x_2 = 1$ and $x_4 = 0$. Solve $R\boldsymbol{x} = \mathbf{0}$. Then $x_1 = -1$ and $x_3 = 0$.

Those numbers -1 and 0 are sitting in column 2 of R (with plus signs).

2. Set $x_2 = 0$ and $x_4 = 1$. Solve $R\boldsymbol{x} = \mathbf{0}$. Then $x_1 = -1$ and $x_3 = -1$.

Those numbers -1 and -1 are sitting in column 4 (with plus signs).

By reversing signs we can read off the special solutions directly from R. The nullspace $N(A) = N(U) = N(R)$ contains all combinations of the special solutions:

$$x = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} = (\textbf{\textit{complete solution of }} Ax = \mathbf{0}).$$

The next section of the book moves firmly from U to R. The MATLAB command $[R, pivcol] = \mathbf{rref}(A)$ produces R and also a list of the pivot columns.

■ REVIEW OF THE KEY IDEAS ■

1. The nullspace $N(A)$, a subspace of $\mathbf{R}^n$, contains all solutions to $Ax = \mathbf{0}$.
2. Elimination produces an echelon matrix U, and then a row reduced R, with pivot columns and free columns.
3. Every free column of U or R leads to a special solution. The free variable equals 1 and the other free variables equal 0. Back substitution solves $Ax = \mathbf{0}$.
4. The complete solution to $Ax = \mathbf{0}$ is a combination of the special solutions.
5. If $n > m$ then A has at least one column without pivots, giving a special solution. So there are nonzero vectors x in the nullspace of this rectangular A.

■ WORKED EXAMPLES ■

3.2 A Create a 3 by 4 matrix whose special solutions to $Ax = \mathbf{0}$ are s_1 and s_2:

$$s_1 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad s_2 = \begin{bmatrix} -2 \\ 0 \\ -6 \\ 1 \end{bmatrix} \qquad \begin{array}{l} \text{pivot columns 1 and 3} \\ \text{free variables } x_2 \text{ and } x_4 \end{array}$$

You could create the matrix A in row reduced form R. Then describe all possible matrices A with the required nullspace $N(A) =$ all combinations of s_1 and s_2.

Solution The reduced matrix R has pivots $= 1$ in columns 1 and 3. There is no third pivot, so the third row of R is all zeros. The free columns 2 and 4 will be combinations of the pivot columns:

$$R = \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{has} \quad Rs_1 = \mathbf{0} \quad \text{and} \quad Rs_2 = \mathbf{0}.$$

The entries $3, 2, 6$ are the negatives of $-3, -2, -6$ in the special solutions!

R is only one matrix (one possible A) with the required nullspace. We could do any elementary operations on R—exchange rows, multiply a row by any $c \neq 0$, subtract any multiple of one row from another. R can be multiplied by *any invertible matrix*, without changing the row space and nullspace.

Every 3 by 4 matrix has at least one special solution. *These A's have two.*

3.2 B Find the special solutions and describe the *complete solution* to $A\boldsymbol{x} = \mathbf{0}$ for

$$A_1 = 3 \text{ by } 4 \textit{ zero matrix} \qquad A_2 = \begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix} \qquad A_3 = \begin{bmatrix} A_2 & A_2 \end{bmatrix}$$

Which are the pivot columns? Which are the free variables? What is R in each case?

Solution $A_1\boldsymbol{x} = \mathbf{0}$ has four special solutions. They are the columns $\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4$ of the 4 by 4 identity matrix. The nullspace is all of $\mathbf{R}^4$. The complete solution is any $\boldsymbol{x} = c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2 + c_3\boldsymbol{s}_3 + c_4\boldsymbol{s}_4$ in $\mathbf{R}^4$. There are no pivot columns; all variables are free; the reduced R is the same zero matrix as A_1.

$A_2\boldsymbol{x} = \mathbf{0}$ has only one special solution $\boldsymbol{s} = (-2, 1)$. The multiples $\boldsymbol{x} = c\boldsymbol{s}$ give the complete solution. The first column of A_2 is its pivot column, and x_2 is the free variable. The row reduced matrices R_2 for A_2 and R_3 for $A_3 = [\, A_2 \;\; A_2 \,]$ have 1's in the pivot:

$$R_2 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \qquad R_3 = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Notice that R_3 has only one pivot column (the first column). All the variables x_2, x_3, x_4 are free. There are three special solutions to $A_3\boldsymbol{x} = \mathbf{0}$ (and also $R_3\boldsymbol{x} = \mathbf{0}$):

$$\boldsymbol{s}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \boldsymbol{s}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \boldsymbol{s}_3 = \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \textbf{Complete solution } \boldsymbol{x} = c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2 + c_3\boldsymbol{s}_3.$$

With r pivots, A has $n - r$ free variables and $A\boldsymbol{x} = \mathbf{0}$ has $n - r$ special solutions.

Problem Set 3.2

Questions 1–4 and 5–8 are about the matrices in Problems 1 and 5.

1 Reduce these matrices to their ordinary echelon forms U:

$$\text{(a)} \quad A = \begin{bmatrix} 1 & 2 & 2 & 4 & 6 \\ 1 & 2 & 3 & 6 & 9 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix} \qquad \text{(b)} \quad B = \begin{bmatrix} 2 & 4 & 2 \\ 0 & 4 & 4 \\ 0 & 8 & 8 \end{bmatrix}.$$

Which are the free variables and which are the pivot variables?

2 For the matrices in Problem 1, find a special solution for each free variable. (Set the free variable to 1. Set the other free variables to zero.)

3 By combining the special solutions in Problem 2, describe every solution to $Ax = \mathbf{0}$ and $Bx = \mathbf{0}$. The nullspace contains only $x = \mathbf{0}$ when there are no ____.

4 By further row operations on each U in Problem 1, find the reduced echelon form R. *True or false*: The nullspace of R equals the nullspace of U.

5 By row operations reduce each matrix to its echelon form U. Write down a 2 by 2 lower triangular L such that $B = LU$.

$$\text{(a)}\quad A = \begin{bmatrix} -1 & 3 & 5 \\ -2 & 6 & 10 \end{bmatrix} \qquad \text{(b)}\quad B = \begin{bmatrix} -1 & 3 & 5 \\ -2 & 6 & 7 \end{bmatrix}.$$

6 Find the special solutions to $Ax = \mathbf{0}$ and $Bx = \mathbf{0}$. For an m by n matrix, the number of pivot variables plus the number of free variables is ____.

7 In Problem 5, describe the nullspaces of A and B in two ways. Give the equations for the plane or the line, and give all vectors x that satisfy those equations as combinations of the special solutions.

8 Reduce the echelon forms U in Problem 5 to R. For each R draw a box around the identity matrix that is in the pivot rows and pivot columns.

Questions 9–17 are about free variables and pivot variables.

9 True or false (with reason if true or example to show it is false):

(a) A square matrix has no free variables.

(b) An invertible matrix has no free variables.

(c) An m by n matrix has no more than n pivot variables.

(d) An m by n matrix has no more than m pivot variables.

10 Construct 3 by 3 matrices A to satisfy these requirements (if possible):

(a) A has no zero entries but $U = I$.

(b) A has no zero entries but $R = I$.

(c) A has no zero entries but $R = U$.

(d) $A = U = 2R$.

11 Put as many 1's as possible in a 4 by 7 echelon matrix U whose pivot variables are

(a) 2, 4, 5

(b) 1, 3, 6, 7

(c) 4 and 6.

12 Put as many 1's as possible in a 4 by 8 *reduced* echelon matrix R so that the free variables are

(a) 2, 4, 5, 6

(b) 1, 3, 6, 7, 8.

13 Suppose column 4 of a 3 by 5 matrix is all zero. Then x_4 is certainly a ____ variable. The special solution for this variable is the vector $\boldsymbol{x} =$ ____.

14 Suppose the first and last columns of a 3 by 5 matrix are the same (not zero). Then ____ is a free variable. Find the special solution for this variable.

15 Suppose an m by n matrix has r pivots. The number of special solutions is ____. The nullspace contains only $\boldsymbol{x} = \mathbf{0}$ when $r =$ ____. The column space is all of $\mathbf{R}^m$ when $r =$ ____.

16 The nullspace of a 5 by 5 matrix contains only $\boldsymbol{x} = \mathbf{0}$ when the matrix has ____ pivots. The column space is $\mathbf{R}^5$ when there are ____ pivots. Explain why.

17 The equation $x - 3y - z = 0$ determines a plane in $\mathbf{R}^3$. What is the matrix A in this equation? Which are the free variables? The special solutions are $(3, 1, 0)$ and ____.

18 (Recommended) The plane $x - 3y - z = 12$ is parallel to the plane $x - 3y - z = 0$ in Problem 17. One particular point on this plane is $(12, 0, 0)$. All points on the plane have the form (fill in the first components)

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} \\ 0 \\ 1 \end{bmatrix}.$$

19 Prove that U and $A = LU$ have the same nullspace when L is invertible:

$$\text{If } U\boldsymbol{x} = \mathbf{0} \text{ then } LU\boldsymbol{x} = \mathbf{0}. \quad \text{If } LU\boldsymbol{x} = \mathbf{0}, \text{ how do you know } U\boldsymbol{x} = \mathbf{0}?$$

20 Suppose column 1 + column 3 + column 5 = $\mathbf{0}$ in a 4 by 5 matrix with four pivots. Which column is sure to have no pivot (and which variable is free)? What is the special solution? What is the nullspace?

Questions 21–28 ask for matrices (if possible) with specific properties.

21 Construct a matrix whose nullspace consists of all combinations of $(2, 2, 1, 0)$ and $(3, 1, 0, 1)$.

22 Construct a matrix whose nullspace consists of all multiples of $(4, 3, 2, 1)$.

23 Construct a matrix whose column space contains $(1, 1, 5)$ and $(0, 3, 1)$ and whose nullspace contains $(1, 1, 2)$.

24 Construct a matrix whose column space contains $(1, 1, 0)$ and $(0, 1, 1)$ and whose nullspace contains $(1, 0, 1)$ and $(0, 0, 1)$.

25 Construct a matrix whose column space contains $(1, 1, 1)$ and whose nullspace is the line of multiples of $(1, 1, 1, 1)$.

26 Construct a 2 by 2 matrix whose nullspace equals its column space. This is possible.

27 Why does no 3 by 3 matrix have a nullspace that equals its column space?

28 If $AB = 0$ then the column space of B is contained in the _____ of A. Give an example of A and B.

29 The reduced form R of a 3 by 3 matrix with randomly chosen entries is almost sure to be _____. What R is virtually certain if the random A is 4 by 3?

30 Show by example that these three statements are generally *false*:

(a) A and A^{T} have the same nullspace.

(b) A and A^{T} have the same free variables.

(c) If R is the reduced form $\mathbf{rref}(A)$ then R^{T} is $\mathbf{rref}(A^{\mathrm{T}})$.

31 If the nullspace of A consists of all multiples of $\boldsymbol{x} = (2, 1, 0, 1)$, how many pivots appear in U? What is R?

32 If the special solutions to $R\boldsymbol{x} = \mathbf{0}$ are in the columns of these N, go backward to find the nonzero rows of the reduced matrices R:

$$N = \begin{bmatrix} 2 & 3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} \\ \\ \end{bmatrix} \text{ (empty 3 by 1).}$$

33 (a) What are the five 2 by 2 reduced echelon matrices R whose entries are all 0's and 1's?

(b) What are the eight 1 by 3 matrices containing only 0's and 1's? Are all eight of them reduced echelon matrices R?

34 Explain why A and $-A$ always have the same reduced echelon form R.

THE RANK AND THE ROW REDUCED FORM ■ 3.3

This section completes the step from A to its reduced row echelon form R. The matrix A is m by n (completely general). The matrix R is also m by n, but each pivot column has only one nonzero entry (the pivot which is always 1). This example is 3 by 5:

$$\textbf{Reduced Row Echelon Form} \qquad R = \begin{bmatrix} \mathbf{1} & 3 & \mathbf{0} & 2 & -1 \\ \mathbf{0} & 0 & \mathbf{1} & 4 & -3 \\ \mathbf{0} & 0 & \mathbf{0} & 0 & 0 \end{bmatrix}.$$

You see zero above the second pivot as well as below. R is the final result of elimination, and MATLAB uses the command **rref**. The Teaching Code **elim** for this book has **rref** built into it. Of course **rref**(R) would give R again!

MATLAB: $[R, pivcol] = \textbf{rref}(A)$ Teaching Code: $[E, R] = \textbf{elim}(A)$

The extra output *pivcol* gives the numbers of the pivot columns. They are the same in A and R. The extra output E is the m by m ***elimination matrix*** that puts the original A (whatever it was) into its row reduced form R:

$$EA = R. \tag{1}$$

The square matrix E is the product of elementary matrices E_{ij} and P_{ij} and D^{-1}. Now we allow $j > i$, when E_{ij} subtracts a multiple of row j from row i. P_{ij} exchanges these rows. D^{-1} divides rows by their pivots to produce 1's.

If we want E, we can apply row reduction to the matrix $\begin{bmatrix} A & I \end{bmatrix}$ with $n + m$ columns. All the elementary matrices that multiply A (to produce R) will also multiply I (to produce E). The whole augmented matrix is being multiplied by E:

$$E\,[A \;\; I] \quad = \quad [R \;\; E] \tag{2}$$

This is exactly what "Gauss-Jordan" did in Chapter 2 to compute A^{-1}. ***When A is square and invertible, its reduced row echelon form is $R = I$.*** Then $EA = R$ becomes $EA = I$. In this invertible case, E is A^{-1}. This chapter is going further, to any (rectangular) matrix A and its reduced form R. The matrix E that multiplies A is still square and invertible, but the best it can do is to produce R. The pivot columns are reduced to ones and zeros.

The Rank of a Matrix

The numbers m and n give the size of a matrix—but not necessarily the *true size* of a linear system. An equation like $0 = 0$ should not count. If there are two identical rows in A, the second one disappears in R. Also if row 3 is a combination of rows 1 and 2, then row 3 will become all zeros in R. We don't want to count rows of zeros. ***The true size of A is given by its rank***:

DEFINITION ***The rank of A is the number of pivots. This number is r.***

The matrix R at the start of this section has rank $r = 2$. It has two pivots and two pivot columns. So does the unknown matrix A that produced R. This number $r = 2$ will be crucial to the theory, but its first definition is entirely computational. To execute the command $r =$ **rank** (A), the computer just counts the pivots. When *pivcol* gives a list of the pivot columns, the length of that list is r.

Actually the computer has a hard time to decide whether a small number is really zero. When it subtracts 3 times $\frac{1}{3}$ from 1, does it obtain zero? Our Teaching Codes treat numbers below the tolerance 10^{-6} as zero.

We know right away that $r \le m$ and $r \le n$. The number of pivots can't be greater than the number of rows. It can't be greater than the number of columns. The cases $r = m$ and $r = n$ of "full row rank" and "full column rank" will be especially important. We mention them here and come back to them soon:

- ***A has full row rank if every row has a pivot: $r = m$. No zero rows in R.***
- ***A has full column rank if every column has a pivot: $r = n$. No free variables.***

A square invertible matrix has $r = m = n$. Then R is the same as I.

At the other extreme are the matrices of ***rank one***. There is only *one pivot*. When elimination clears out the first column, it clears out all the columns. *Every row is a multiple of the pivot row.* At the same time, every column is a multiple of the pivot column!

$$\textbf{Rank one matrix} \qquad A = \begin{bmatrix} \mathbf{1} & 3 & 10 \\ \mathbf{2} & 6 & 20 \\ \mathbf{3} & 9 & 30 \end{bmatrix} \longrightarrow R = \begin{bmatrix} \mathbf{1} & 3 & 10 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The column space of a rank one matrix is "one-dimensional". Here all columns are on the line through $\boldsymbol{u} = (1, 2, 3)$. The columns of A are $\boldsymbol{u}$ and $3\boldsymbol{u}$ and $10\boldsymbol{u}$. Put those numbers into the row $\boldsymbol{v}^{\mathrm{T}} = [\ 1 \quad 3 \quad 10\ ]$ and you have the special rank one form $A = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$:

$$A = \textbf{column times row} = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} \qquad \begin{bmatrix} 1 & 3 & 10 \\ 2 & 6 & 20 \\ 3 & 9 & 30 \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \end{bmatrix} [1 \quad 3 \quad 10]\,. \tag{3}$$

Example 1 When all rows are multiples of one pivot row, the rank is $r = 1$:

$$\begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 8 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 3 \\ 0 & 5 \end{bmatrix} \text{ and } \begin{bmatrix} 5 \\ 2 \end{bmatrix} \text{ and } \begin{bmatrix} 6 \end{bmatrix} \text{ all have rank 1.}$$

The reduced row echelon forms $R =$ **rref** (A) can be checked by eye:

$$R = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \end{bmatrix} \text{ have only one pivot.}$$

Our second definition of rank is coming at a higher level. It deals with entire rows and entire columns—vectors and not just numbers. The matrices A and U and R

have r *independent* rows (the pivot rows). They also have r *independent* columns (the pivot columns). Section 3.5 says what it means for rows or columns to be independent.

A third definition of rank, at the top level of linear algebra, will deal with *spaces* of vectors. The rank r is the "dimension" of the column space. It is also the dimension of the row space. The great thing is that r also reveals the dimension of the nullspace.

The Pivot Columns

The pivot columns of R have 1's in the pivots and 0's everywhere else. The r pivot columns taken together contain an r by r identity matrix I. It sits above $m - r$ rows of zeros. The numbers of the pivot columns are in the list *pivcol*.

The pivot columns of A are probably *not* obvious from A itself. But their column numbers are given by the *same list pivcol*. The r columns of A that eventually have pivots (in U and R) are the pivot columns. The first matrix R in this section is the row reduced echelon form of this matrix A, with $pivcol = (1, 3)$:

Pivot Columns
$$A = \begin{bmatrix} \mathbf{1} & 3 & \mathbf{0} & 2 & -1 \\ \mathbf{0} & 0 & \mathbf{1} & 4 & -3 \\ \mathbf{1} & 3 & \mathbf{1} & 6 & -4 \end{bmatrix} \quad \text{yields } R = \begin{bmatrix} \mathbf{1} & 3 & \mathbf{0} & 2 & -1 \\ \mathbf{0} & 0 & \mathbf{1} & 4 & -3 \\ \mathbf{0} & 0 & \mathbf{0} & 0 & 0 \end{bmatrix}.$$

The column spaces of R and A can be different! All columns of this R end with zeros. E subtracts rows 1 and 2 of A from row 3 (to produce that zero row in R):

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \quad \text{and} \quad E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

The r pivot columns of A are also the first r columns of E^{-1}. The reason is that each column of A is E^{-1} times a column of R. The r by r identity matrix inside R just picks out the first r columns of E^{-1}.

One more fact about pivot columns. Their definition has been purely computational, based on R. Here is a direct mathematical description of the pivot columns of A:

3C ***The pivot columns are not combinations of earlier columns.*** The free columns ***are*** combinations of earlier columns. These combinations are the special solutions!

A pivot column of R (with 1 in the pivot row) cannot be a combination of earlier columns (with 0's in that row). The same column of A can't be a combination of earlier columns, because $A\boldsymbol{x} = \mathbf{0}$ *exactly when* $R\boldsymbol{x} = \mathbf{0}$. Now we look at the special solution $\boldsymbol{x}$ from each free column.

The Special Solutions

Each special solution to $A\boldsymbol{x} = \mathbf{0}$ and $R\boldsymbol{x} = \mathbf{0}$ has one free variable equal to 1. The other free variables are all zero. The solutions come directly from the echelon form R:

Free columns
Free variables

$$R\boldsymbol{x} = \begin{bmatrix} 1 & \mathbf{3} & 0 & \mathbf{2} & \mathbf{-1} \\ 0 & \mathbf{0} & 1 & \mathbf{4} & \mathbf{-3} \\ 0 & \mathbf{0} & 0 & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} x_1 \\ \boldsymbol{x_2} \\ x_3 \\ \boldsymbol{x_4} \\ \boldsymbol{x_5} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The free variables are in boldface. Set the first free variable to $x_2 = 1$ with $x_4 = x_5 = 0$. The equations give the pivot variables $x_1 = -3$ and $x_3 = 0$. This says that column 2 (a free column) is 3 times column 1. The special solution is $\boldsymbol{s}_1 = (-3, 1, 0, 0, 0)$.

The next special solution has $x_4 = 1$. The other free variables are $x_2 = x_5 = 0$. The solution is $\boldsymbol{s}_2 = (-2, 0, -4, 1, 0)$. Notice -2 and -4 in R, with plus signs.

The third special solution has $x_5 = 1$. With $x_2 = 0$ and $x_4 = 0$ we find $\boldsymbol{s}_3 = (1, 0, 3, 0, 1)$. The numbers $x_1 = 1$ and $x_3 = 3$ are in column 5 of R, again with opposite signs. This is a general rule as we soon verify. The nullspace matrix N contains the three special solutions in its columns:

Nullspace matrix

$$N = \begin{array}{c} \begin{bmatrix} \mathbf{-3} & \mathbf{-2} & \mathbf{1} \\ 1 & 0 & 0 \\ \mathbf{0} & \mathbf{-4} & \mathbf{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{l} \text{not free} \\ \text{free} \\ \text{not free} \\ \text{free} \\ \text{free} \end{array}$$

The linear combinations of these three columns give all vectors in the nullspace. This is the complete solution to $A\boldsymbol{x} = \mathbf{0}$ (and $R\boldsymbol{x} = \mathbf{0}$). Where R had the identity matrix (2 by 2) in its pivot columns, N has the identity matrix (3 by 3) in its free rows.

There is a special solution for every free variable. Since r columns have pivots, that leaves $n - r$ free variables. This is the key to $A\boldsymbol{x} = \mathbf{0}$.

3D $A\boldsymbol{x} = \mathbf{0}$ ***has*** $n - r$ ***free variables and special solutions***: n columns minus r pivot columns. The ***nullspace matrix*** N has $n - r$ columns (the special solutions).

When we introduce the idea of "independent" vectors, we will show that the special solutions are independent. You can see in N that no column is a combination of the other columns. The beautiful thing is that the count is exactly right:

*$A\boldsymbol{x} = \mathbf{0}$ **has r independent equations so $n - r$ independent solutions.***

To complete this section, look again at the special solutions. Suppose for simplicity that the first r columns are the pivot columns, and the last $n-r$ columns are free (no pivots). Then the reduced row echelon form looks like

$$R = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{F} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{matrix} r \text{ pivot rows} \\ m-r \text{ zero rows} \end{matrix} \tag{4}$$

r pivot columns $\quad$ $n-r$ free columns

3E The pivot variables in the $n-r$ special solutions come by changing F to $-F$:

$$\textbf{Nullspace matrix} \quad N = \begin{bmatrix} -\boldsymbol{F} \\ \boldsymbol{I} \end{bmatrix} \begin{matrix} r \text{ pivot variables} \\ n-r \text{ free variables} \end{matrix} \tag{5}$$

Check $RN = 0$. The first block row of RN is (I times $-F$) + (F times I) = zero. The columns of N solve $R\boldsymbol{x} = \mathbf{0}$. When the free part of $R\boldsymbol{x} = \mathbf{0}$ moves to the right side, the left side just holds the identity matrix:

$$I \begin{bmatrix} \text{pivot} \\ \text{variables} \end{bmatrix} = -F \begin{bmatrix} \text{free} \\ \text{variables} \end{bmatrix}. \tag{6}$$

In each special solution, the free variables are a column of I. Then the pivot variables are a column of $-F$. Those special solutions give the nullspace matrix N.

The idea is still true if the pivot columns are mixed in with the free columns. Then I and F are mixed together. You can still see $-F$ in the solutions. Here is an example where $I = [1]$ comes first and $F = \begin{bmatrix} 2 & 3 \end{bmatrix}$ comes last.

Example 2 The special solutions of $R\boldsymbol{x} = x_1 + 2x_2 + 3x_3 = 0$ are the columns of N:

$$R = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \qquad N = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The rank is one. There are $n-r = 3-1$ special solutions $(-2, 1, 0)$ and $(-3, 0, 1)$.

Final Note How can I write confidently about R when I don't know which steps MATLAB will take? A could be reduced to R in different ways. Very likely you and Mathematica and Maple would do the elimination differently. The key point is that ***the final matrix R is always the same***. The original A *completely determines the* I *and* F *and zero rows in* R, according to **3C**:

The pivot columns *are not* combinations of earlier columns of A.

The free columns *are* combinations of earlier columns (F tells the combinations).

A small example with rank one will show two E's that produce the correct $EA = R$:

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{reduces to} \quad R = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and no other } R.$$

You could multiply row 1 of A by $\frac{1}{2}$, and subtract row 1 from row 2:

$$\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ -1/2 & 1 \end{bmatrix} = E.$$

Or you could exchange rows in A, and then subtract 2 times row 1 from row 2:

$$\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} = E_{\text{new}}.$$

Multiplication gives $EA = R$ and also $E_{\text{new}}A = R$. ***Different E's but the same R.***

■ REVIEW OF THE KEY IDEAS ■

1. The rank of A is the number of pivots (which are 1's in R).

2. The r pivot columns of A and R are in the same list *pivcol*.

3. Those r pivot columns are not combinations of earlier columns.

4. The $n - r$ free columns are combinations of earlier columns.

5. Those combinations (using $-F$ taken from R) give the $n-r$ special solutions to $A\boldsymbol{x} = \boldsymbol{0}$ and $R\boldsymbol{x} = \boldsymbol{0}$. They are the $n - r$ columns of the nullspace matrix N.

■ WORKED EXAMPLES ■

3.3 A Factor these rank one matrices into $A = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ = row times column:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} \qquad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad (\text{find } d \text{ from } a^{-1}, b, c)$$

Split this rank two matrix into $\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}} + \boldsymbol{u}_2\boldsymbol{v}_2^{\mathrm{T}}$ = (3 by 2) times (2 by 4) using E^{-1} and R:

$$A = \begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 0 & 3 \\ 2 & 3 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = E^{-1}R.$$

Solution For the 3 by 3 matrix A, all rows are multiples of $\boldsymbol{v}^{\mathrm{T}} = [1 \ \ 2 \ \ 3]$. All columns are multiples of the column $\boldsymbol{u} = (1, 2, 3)$. This symmetric matrix has $\boldsymbol{u} = \boldsymbol{v}$ and A is $\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$. Every rank one symmetric matrix will have this form or else $-\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$.

If the 2 by 2 matrix $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ has rank one, it must be singular. In Chapter 5, its determinant is $ad - bc = 0$. In this chapter, row 2 is a multiple of row 1. That multiple is $\frac{c}{a}$ (the problem assumes $a \neq 0$). Rank one always produces column times row:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 \\ c/a \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} a & b \\ c & bc/a \end{bmatrix}. \quad \text{So} \quad d = \frac{bc}{a}.$$

The 3 by 4 matrix of rank two is a sum of *two matrices of rank one*. All columns of A are combinations of the pivot columns 1 and 2. All rows are combinations of the nonzero rows of R. The pivot columns are $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ and those nonzero rows are $\boldsymbol{v}_1^{\mathrm{T}}$ and $\boldsymbol{v}_2^{\mathrm{T}}$. Then A is $\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}} + \boldsymbol{u}_2\boldsymbol{v}_2^{\mathrm{T}}$, multiplying columns of E^{-1} times rows of R:

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 0 & 3 \\ 2 & 3 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}$$

3.3 B Find the row reduced form R and the rank r of A (*those depend on* c). Which are the pivot columns of A? Which variables are free? What are the special solutions and the nullspace matrix N (always depending on c)?

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 3 \\ 4 & 8 & c \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} c & c \\ c & c \end{bmatrix}.$$

Solution The 3 by 3 matrix A has rank $r = 2$ ***except if*** $c = 4$. The pivots are in columns 1 and 3. The second variable x_2 is free. Notice the form of R:

$$c \neq 4 \quad R = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad c = 4 \quad R = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

When $c = 4$, the only pivot is in column 1 (one pivot column). Columns 2 and 3 are multiples of column 1 (so rank $= 1$). The second and third variables are free, producing two special solutions:

$$c \neq 4 \quad \text{Special solution with } x_2 = 1 \text{ goes into} \quad N = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}.$$

$$c = 4 \quad \text{Another special solution goes into} \quad N = \begin{bmatrix} -2 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The 2 by 2 matrix $\left[\begin{smallmatrix} c & c \\ c & c \end{smallmatrix}\right]$ has rank $r = 1$ ***except if*** $c = 0$, when the rank is zero!

$$c \neq 0 \quad R = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad\qquad c = 0 \quad R = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The first column is the pivot column if $c \neq 0$, and the second variable is free (one special solution in N). The matrix has *no pivot columns* if $c = 0$, and both variables are free:

$$c \neq 0 \quad N = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad\qquad c = 0 \quad N = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Problem Set 3.3

1 Which of these rules gives a correct definition of the *rank* of A?

(a) The number of nonzero rows in R.

(b) The number of columns minus the total number of rows.

(c) The number of columns minus the number of free columns.

(d) The number of 1's in the matrix R.

2 Find the reduced row echelon forms R and the rank of these matrices:

(a) The 3 by 4 matrix of all ones.

(b) The 3 by 4 matrix with $a_{ij} = i + j - 1$.

(c) The 3 by 4 matrix with $a_{ij} = (-1)^j$.

3 Find R for each of these (block) matrices:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 3 \\ 2 & 4 & 6 \end{bmatrix} \quad B = \begin{bmatrix} A & A \end{bmatrix} \quad C = \begin{bmatrix} A & A \\ A & 0 \end{bmatrix}$$

4 Suppose all the pivot variables come *last* instead of first. Describe all four blocks in the reduced echelon form (the block B should be r by r):

$$R = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

What is the nullspace matrix N containing the special solutions?

5 (Silly problem) Describe all 2 by 3 matrices A_1 and A_2, with row echelon forms R_1 and R_2, such that $R_1 + R_2$ is the row echelon form of $A_1 + A_2$. Is is true that $R_1 = A_1$ and $R_2 = A_2$ in this case?

6 If A has r pivot columns, how do you know that A^{T} has r pivot columns? Give a 3 by 3 example for which the column numbers are different.

7 What are the special solutions to $Rx = \mathbf{0}$ and $y^{\mathrm{T}}R = \mathbf{0}$ for these R?

$$R = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad R = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Problems 8–11 are about matrices of rank $r = 1$.

8 Fill out these matrices so that they have rank 1:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & & \\ 4 & & \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & & \\ 1 & & \\ 2 & 6 & -3 \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} a & b \\ c & \end{bmatrix}.$$

9 If A is an m by n matrix with $r = 1$, its columns are multiples of one column and its rows are multiples of one row. The column space is a _____ in $\mathbf{R}^m$. The nullspace is a _____ in $\mathbf{R}^n$. Also the column space of A^{T} is a _____ in $\mathbf{R}^n$.

10 Choose vectors u and v so that $A = uv^{\mathrm{T}}$ = column times row:

$$A = \begin{bmatrix} 3 & 6 & 6 \\ 1 & 2 & 2 \\ 4 & 8 & 8 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 2 & 2 & 6 & 4 \\ -1 & -1 & -3 & -2 \end{bmatrix}.$$

$A = uv^{\mathrm{T}}$ is the natural form for every matrix that has rank $r = 1$.

11 If A is a rank one matrix, the second row of U is _____. Do an example.

Problems 12–14 are about r by r invertible matrices inside A.

12 *If A has rank r, then it has an r by r submatrix S that is invertible.* Remove $m - r$ rows and $n - r$ columns to find an invertible submatrix S inside each A (you could keep the pivot rows and pivot columns of A):

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

13 Suppose P is the submatrix of A containing only the pivot columns. Explain why this m by r submatrix P has rank r.

14 In Problem 13, we can transpose P and find the r pivot columns of P^{T}. Transposing back, we have an r by r invertible submatrix S inside P:

$$\text{For } A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 4 & 7 \end{bmatrix} \text{ find } P \text{ (3 by 2) and then } S \text{ (2 by 2).}$$

Problems 15–20 show that rank(AB) is not greater than rank(A) or rank(B).

15 Find the ranks of AB and AM (rank one matrix times rank one matrix):

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 1 & 4 \\ 3 & 1.5 & 6 \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} 1 & b \\ c & bc \end{bmatrix}.$$

16 The rank one matrix $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ times the rank one matrix $\boldsymbol{w}\boldsymbol{z}^{\mathrm{T}}$ is $\boldsymbol{u}\boldsymbol{z}^{\mathrm{T}}$ times the number ____. This has rank one unless ____ $= 0$.

17 (a) Suppose column j of B is a combination of previous columns of B. Show that column j of AB is the same combination of previous columns of AB. Then $\mathbf{rank}(\boldsymbol{AB}) \leq \mathbf{rank}(\boldsymbol{B})$, because AB cannot have new pivot columns.

(b) Find A_1 and A_2 so that rank$(A_1B) = 1$ and rank$(A_2B) = 0$ for $B = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$.

18 Problem 17 proved that rank$(AB) \leq$ rank(B). Then the same reasoning gives rank$(B^{\mathrm{T}}A^{\mathrm{T}}) \leq$ rank(A^{T}). How do you deduce that $\mathbf{rank}(\boldsymbol{AB}) \leq \mathbf{rank}\, \boldsymbol{A}$?

19 *(Important)* Suppose A and B are n by n matrices, and $AB = I$. Prove from rank$(AB) \leq$ rank(A) that the rank of A is n. So A is invertible and B must be its two-sided inverse (Section 2.5). Therefore $BA = I$ *(which is not so obvious!)*.

20 If A is 2 by 3 and B is 3 by 2 and $AB = I$, show from its rank that $BA \neq I$. Give an example of A and B. For $m < n$, a right inverse is not a left inverse.

21 Suppose A and B have the *same* reduced row echelon form R.

(a) Show that A and B have the same nullspace and the same row space.

(b) We know $E_1A = R$ and $E_2B = R$. So A equals an ____ matrix times B.

22 Every m by n matrix of rank r reduces to (m by r) times (r by n):

$$A = (\text{pivot columns of } A)\ (\text{first } r \text{ rows of } R) = (\mathbf{COL})(\mathbf{ROW})^{\mathrm{T}}.$$

Write the 3 by 5 matrix A at the start of this section as the product of the 3 by 2 matrix from the pivot columns and the 2 by 5 matrix from R.

23 $A = (\mathrm{COL})(\mathrm{ROW})^{\mathrm{T}}$ is a sum of r rank one matrices (multiply columns times rows). Express A and B as the sum of two rank one matrices:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 4 \\ 1 & 1 & 8 \end{bmatrix} \qquad B = \begin{bmatrix} A & A \end{bmatrix}.$$

24 Suppose A is an m by n matrix of rank r. Its reduced echelon form is R. Describe exactly the matrix Z (its shape and all its entries) that comes from *transposing the reduced row echelon form of* R' (prime means transpose):

$$R = \mathbf{rref}(A) \quad \text{and} \quad Z = (\mathbf{rref}(R'))'.$$

25 Instead of transposing R (Problem 24) we could transpose A first. Explain in one line why $Y = Z$:

$$V = \mathbf{rref}(A') \quad \text{and} \quad Y = \mathbf{rref}(V').$$

26 Answer the same questions as in Worked Example **3.3 B** for

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 4 & 4 \\ 1 & c & 2 & 2 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1-c & 2 \\ 0 & 2-c \end{bmatrix}.$$

27 What is the nullspace matrix N (containing the special solutions) for A, B, C?

$$A = \begin{bmatrix} I & I \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} I & I \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad C = [\, I \;\; I \;\; I \,].$$

THE COMPLETE SOLUTION TO $AX = B$ ■ 3.4

The last section totally solved $A\boldsymbol{x} = \mathbf{0}$. Elimination converted the problem to $R\boldsymbol{x} = \mathbf{0}$. The free variables were given special values (one and zero). Then the pivot variables were found by back substitution. We paid no attention to the right side $\boldsymbol{b}$ because it started and ended as zero. The solution $\boldsymbol{x}$ was in the nullspace of A.

Now $\boldsymbol{b}$ is not zero. Row operations on the left side must act also on the right side. One way to organize that is to ***add b as an extra column of the matrix***. We keep the same example A as before. But we *"augment"* A with the right side $(b_1, b_2, b_3) = (1, 6, 7)$:

$$\begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 1 & 3 & 1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 7 \end{bmatrix} \quad \begin{matrix} \text{has the} \\ \text{augmented} \\ \text{matrix} \end{matrix} \quad \begin{bmatrix} 1 & 3 & 0 & 2 & \mathbf{1} \\ 0 & 0 & 1 & 4 & \mathbf{6} \\ 1 & 3 & 1 & 6 & \mathbf{7} \end{bmatrix} = \begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}.$$

The augmented matrix is just $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$. When we apply the usual elimination steps to A, we also apply them to $\boldsymbol{b}$. In this example we subtract row 1 from row 3 and then subtract row 2 from row 3. This produces a *complete row of zeros*:

$$\begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix} \quad \begin{matrix} \text{has the} \\ \text{augmented} \\ \text{matrix} \end{matrix} \quad \begin{bmatrix} 1 & 3 & 0 & 2 & \mathbf{1} \\ 0 & 0 & 1 & 4 & \mathbf{6} \\ 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix} = \begin{bmatrix} R & \boldsymbol{d} \end{bmatrix}.$$

That very last zero is crucial. It means that the equations can be solved; the third equation has become $0 = 0$. In the original matrix A, the first row plus the second row equals the third row. If the equations are consistent, this must be true on the right side of the equations also! The all-important property on the right side was $1 + 6 = 7$.

Here are the same augmented matrices for a general $\boldsymbol{b} = (b_1, b_2, b_3)$:

$$\begin{bmatrix} 1 & 3 & 0 & 2 & b_1 \\ 0 & 0 & 1 & 4 & b_2 \\ 1 & 3 & 1 & 6 & b_3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 3 & 0 & 2 & b_1 \\ 0 & 0 & 1 & 4 & b_2 \\ 0 & 0 & 0 & 0 & b_3 - b_1 - b_2 \end{bmatrix}.$$

Now we get $0 = 0$ in the third equation provided $b_3 - b_1 - b_2 = 0$. This is $b_1 + b_2 = b_3$.

One Particular Solution

Choose the free variables to be $x_2 = x_4 = 0$. Then the equations give the pivot variables $x_1 = 1$ and $x_3 = 6$. They are in the last column $\boldsymbol{d}$ of the reduced augmented matrix. The code $\boldsymbol{x} =$ **partic** $(A, \boldsymbol{b})$ gives this particular solution (call it $\boldsymbol{x}_p$) to $A\boldsymbol{x} = \boldsymbol{b}$. First A and $\boldsymbol{b}$ reduce to R and $\boldsymbol{d}$. Zero rows in R must also be zero in $\boldsymbol{d}$. Then the r pivot variables in $\boldsymbol{x}$ are taken directly from $\boldsymbol{d}$, because the pivot columns in R contain the identity matrix. After row reduction we are just solving $I\boldsymbol{x} = \boldsymbol{d}$.

Notice how we *choose* the free variables (as zero) and *solve* for the pivot variables. After the row reduction to R, those steps are quick. When the free variables are zero, the pivot variables for $\boldsymbol{x}_p$ are in the extra column:

The particular solution solves $\quad A\boldsymbol{x}_p = \boldsymbol{b}$
The $n - r$ special solutions solve $\quad A\boldsymbol{x}_n = \boldsymbol{0}$.

In this example the particular solution is $(1, 0, 6, 0)$. The two special (nullspace) solutions to $R\boldsymbol{x} = \boldsymbol{0}$ come from the two free columns of R, by reversing signs of 3, 2, and 4. ***Please notice how I write the complete solution*** $\boldsymbol{x}_p + \boldsymbol{x}_n$ ***to*** $A\boldsymbol{x} = \boldsymbol{b}$:

$$\textbf{Complete solution:} \quad \boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n = \begin{bmatrix} 1 \\ 0 \\ 6 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ -4 \\ 1 \end{bmatrix}.$$

Question Suppose A is a square invertible matrix, $m = n = r$. What are $\boldsymbol{x}_p$ and $\boldsymbol{x}_n$?

Answer The particular solution is the one and *only* solution $A^{-1}\boldsymbol{b}$. There are no special solutions or free variables. $R = I$ has no zero rows. The only vector in the nullspace is $\boldsymbol{x}_n = \boldsymbol{0}$. The complete solution is $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n = A^{-1}\boldsymbol{b} + \boldsymbol{0}$.

This was the situation in Chapter 2. We didn't mention the nullspace in that chapter. $N(A)$ contained only the zero vector. Reduction goes from $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$ to $\begin{bmatrix} I & A^{-1}\boldsymbol{b} \end{bmatrix}$. The original $A\boldsymbol{x} = \boldsymbol{b}$ is reduced all the way to $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. This is a special case here, but square invertible matrices are the ones we see most often in practice. So they got their own chapter at the start of the book.

For small examples we can put $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$ into reduced row echelon form. For a large matrix, MATLAB can do it better. Here is a small example with *full column rank*. Both columns have pivots.

Example 1 Find the condition on (b_1, b_2, b_3) for $A\boldsymbol{x} = \boldsymbol{b}$ to be solvable, if

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ -2 & -3 \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

This condition puts $\boldsymbol{b}$ in the column space of A. Find the complete $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n$.
Solution Use the augmented matrix, with its extra column $\boldsymbol{b}$. Elimination subtracts row 1 from row 2, and adds 2 times row 1 to row 3:

$$\begin{bmatrix} 1 & 1 & b_1 \\ 1 & 2 & b_2 \\ -2 & -3 & b_3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & b_1 \\ 0 & 1 & b_2 - b_1 \\ 0 & -1 & b_3 + 2b_1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2b_1 - b_2 \\ 0 & 1 & b_2 - b_1 \\ 0 & 0 & b_3 + b_1 + b_2 \end{bmatrix}.$$

The last equation is $0 = 0$ provided $b_3 + b_1 + b_2 = 0$. This is the condition to put $\boldsymbol{b}$ in the column space; then the system is solvable. The rows of A add to the zero row. So for consistency (these are equations!) the entries of $\boldsymbol{b}$ must also add to zero.

This example has no free variables and no special solutions. The nullspace solution is $\boldsymbol{x}_n = \boldsymbol{0}$. The (only) particular solution $\boldsymbol{x}_p$ is at the top of the augmented column:

$$\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n = \begin{bmatrix} 2b_1 - b_2 \\ b_2 - b_1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

If $b_3 + b_1 + b_2$ is not zero, there is no solution to $A\boldsymbol{x} = \boldsymbol{b}$ ($\boldsymbol{x}_p$ doesn't exist).

This example is typical of the extremely important case when A has *full column rank*. Every column has a pivot. *The rank is* $r = n$. The matrix is tall and thin ($m \geq n$). Row reduction puts I at the top, when A is reduced to R:

$$\textbf{Full column rank} \quad R = \begin{bmatrix} n \text{ by } n \text{ identity matrix} \\ m - n \text{ rows of zeros} \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix}. \tag{1}$$

There are no free columns or free variables. The nullspace matrix is empty!

We will collect together the different ways of recognizing this type of matrix.

3F Every matrix A with **full column rank** ($r = n$) has all these properties:

1. All columns of A are pivot columns.
2. There are no free variables or special solutions.
3. The nullspace $N(A)$ contains only the zero vector $\boldsymbol{x} = \boldsymbol{0}$.
4. If $A\boldsymbol{x} = \boldsymbol{b}$ has a solution (it might not) then it has only *one solution*.

In the language of the next section, this A has ***independent columns***. In Chapter 4 we will add one more fact to the list: *The square matrix $A^{\mathrm{T}}A$ is invertible.*

In this case the nullspace of A (and R) has shrunk to the zero vector. The solution to $A\boldsymbol{x} = \boldsymbol{b}$ is *unique* (if it exists). There will be $m - n$ (here $3 - 2$) zero rows in R. So there are $m - n$ (here 1 condition) conditions on $\boldsymbol{b}$ in order to have $0 = 0$ in those rows. If $b_3 + b_1 + b_2 = 0$ is satisfied, $A\boldsymbol{x} = \boldsymbol{b}$ has exactly one solution.

The Complete Solution

The other extreme case is full row rank. Now $A\boldsymbol{x} = \boldsymbol{b}$ either has one or infinitely may solutions. In this case A is *short and wide* ($m \leq n$). The number of unknowns is at least the number of equations. A matrix has ***full row rank*** if $r = m$. The nullspace of A^{T} shrinks to the zero vector. Every row has a pivot, and here is an example.

Example 2 There are $n = 3$ unknowns but only two equations. The rank is $r = m = 2$:

$$\begin{aligned} x \;+\; y \;+\; z &= 3 \\ x \;+\; 2y \;-\; z &= 4 \end{aligned}$$

These are two planes in xyz space. The planes are not parallel so they intersect in a line. This line of solutions is exactly what elimination will find. ***The particular solution will be one point on the line. Adding the nullspace vectors $\boldsymbol{x}_n$ will move us along the line.*** Then $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n$ gives the whole line of solutions.

We find $\boldsymbol{x}_p$ and $\boldsymbol{x}_n$ by elimination. Subtract row 1 from row 2 and then subtract row 2 from row 1:

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{3} \\ 1 & 2 & -1 & \mathbf{4} \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & \mathbf{3} \\ 0 & 1 & -2 & \mathbf{1} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 3 & \mathbf{2} \\ 0 & 1 & -2 & \mathbf{1} \end{bmatrix} = \begin{bmatrix} R & \boldsymbol{d} \end{bmatrix}.$$

The particular solution has free variable $x_3 = 0$. The special solution has $x_3 = 1$:

$\boldsymbol{x}_{\textbf{particular}}$ comes directly from $\boldsymbol{d}$ the right side: $\boldsymbol{x}_p = (2, 1, 0)$
$\boldsymbol{x}_{\textbf{special}}$ comes from the third column (free column F) of R: $\boldsymbol{s} = (-3, 2, 1)$

It is wise to check that $\boldsymbol{x}_p$ and $\boldsymbol{s}$ satisfy the original equations $A\boldsymbol{x}_p = \boldsymbol{b}$ and $A\boldsymbol{s} = \boldsymbol{0}$:

$$\begin{aligned} 2 + 1 &= 3 \qquad & -3 + 2 + 1 &= 0 \\ 2 + 2 &= 4 \qquad & -3 + 4 - 1 &= 0 \end{aligned}$$

The nullspace solution $\boldsymbol{x}_n$ is any multiple of $\boldsymbol{s}$. It moves along the line of solutions, starting at $\boldsymbol{x}_{\text{particular}}$. ***Please notice again how to write the answer***:

$$\textbf{Complete Solution:} \qquad \boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}.$$

This line is drawn in Figure 3.3. Any point on the line *could* have been chosen as the particular solution; we chose the point with $x_3 = 0$. The particular solution is *not* multiplied by an arbitrary constant! The special solution is, and you understand why.

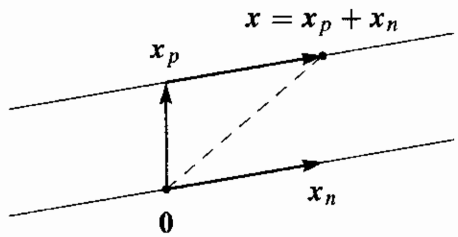


Figure 3.3 The complete solution is *one* particular solution plus *all* nullspace solutions.

Now we summarize this short wide case ($m \le n$) of *full row rank*:

3G Every matrix A with ***full row rank*** ($r = m$) has all these properties:

1. All rows have pivots, and R has no zero rows.
2. $Ax = b$ has a solution for every right side b.
3. The column space is the whole space $\mathbf{R}^m$.
4. There are $n - r = n - m$ special solutions in the nullspace of A.

In this case with m pivots, the rows are "linearly independent". In other words, the columns of A^{T} are linearly independent. We are more than ready for the definition of linear independence, as soon as we summarize the four possibilities—which depend on the rank. Notice how r, m, n are the critical numbers!

The four possibilities for linear equations depend on the rank r:

$r = m$	and	$r = n$	*Square and invertible*	$Ax = b$	has 1 solution
$r = m$	and	$r < n$	*Short and wide*	$Ax = b$	has ∞ solutions
$r < m$	and	$r = n$	*Tall and thin*	$Ax = b$	has 0 or 1 solution
$r < m$	and	$r < n$	*Unknown shape*	$Ax = b$	has 0 or ∞ solutions

The reduced R will fall in the same category as the matrix A. In case the pivot columns happen to come first, we can display these four possibilities for R:

$$R = \begin{bmatrix} I \end{bmatrix} \qquad \begin{bmatrix} I & F \end{bmatrix} \qquad \begin{bmatrix} I \\ 0 \end{bmatrix} \qquad \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$$
$$r = m = n \qquad r = m < n \qquad r = n < m \qquad r < m, r < n$$

Cases 1 and 2 have full row rank $r = m$. Cases 1 and 3 have full column rank $r = n$. Case 4 is the most general in theory and the least common in practice.

Note In the first edition of this textbook, we generally stopped at U before reaching R. Instead of reading the complete solution directly from $R\boldsymbol{x} = \boldsymbol{d}$, we found it by back substitution from $U\boldsymbol{x} = \boldsymbol{c}$. That combination of reduction to U and back substitution for $\boldsymbol{x}$ is slightly faster. Now we prefer the complete reduction: a single "1" in each pivot column. We find that everything is so much clearer in R (and the computer should do the hard work anyway) that we reduce all the way.

■ REVIEW OF THE KEY IDEAS ■

1. The rank r is the number of pivots. The matrix R has $m - r$ zero rows.
2. $A\boldsymbol{x} = \boldsymbol{b}$ is solvable if and only if the last $m - r$ equations reduce to $0 = 0$.
3. One particular solution $\boldsymbol{x}_p$ has all free variables equal to zero.
4. The pivot variables are determined after the free variables are chosen.
5. Full column rank $r = n$ means no free variables: one solution or none.
6. Full row rank $r = m$ means one solution if $m = n$ or infinitely many if $m < n$.

■ WORKED EXAMPLES ■

3.4 A This question connects elimination-pivot columns-back substitution to column space-nullspace-rank-solvability (the full picture). The 3 by 3 matrix A has rank 2:

$$A\boldsymbol{x} = \boldsymbol{b} \quad \text{is} \quad \begin{aligned} x_1 + 2x_2 + 3x_3 + \ \ 5x_4 &= b_1 \\ 2x_1 + 4x_2 + 8x_3 + 12x_4 &= b_2 \\ 3x_1 + 6x_2 + 7x_3 + 13x_4 &= b_3 \end{aligned}$$

1. Reduce $[\,A \ \ \boldsymbol{b}\,]$ to $[\,U \ \ \boldsymbol{c}\,]$, so that $A\boldsymbol{x} = \boldsymbol{b}$ becomes a triangular system $U\boldsymbol{x} = \boldsymbol{c}$.
2. Find the condition on b_1, b_2, b_3 for $A\boldsymbol{x} = \boldsymbol{b}$ to have a solution.
3. Describe the column space of A. Which plane in $\mathbf{R}^3$?
4. Describe the nullspace of A. Which special solutions in $\mathbf{R}^4$?
5. Find a particular solution to $A\boldsymbol{x} = (0, 6, -6)$ and then the complete solution.
6. Reduce $[\,U \ \ c\,]$ to $[\,R \ \ \boldsymbol{d}\,]$: Special solutions from R, particular solution from $\boldsymbol{d}$.

Solution

1. The multipliers in elimination are 2 and 3 and -1. They take $[\,A \ \ \boldsymbol{b}\,]$ into $[\,U \ \ \boldsymbol{c}\,]$.

$$\begin{bmatrix} 1 & 2 & 3 & 5 & \boldsymbol{b_1} \\ 2 & 4 & 8 & 12 & \boldsymbol{b_2} \\ 3 & 6 & 7 & 13 & \boldsymbol{b_3} \end{bmatrix} \rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 3 & 5 & \boldsymbol{b_1} \\ 0 & 0 & 2 & 2 & \boldsymbol{b_2 - 2b_1} \\ 0 & 0 & -2 & -2 & \boldsymbol{b_3 - 3b_1} \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 3 & 5 & \boldsymbol{b_1} \\ 0 & 0 & 2 & 2 & \boldsymbol{b_2 - 2b_1} \\ 0 & 0 & 0 & 0 & \boldsymbol{b_3 + b_2 - 5b_1} \end{array}\right]$$

2. The last equation shows the solvability condition $b_3+b_2-5b_1=0$. Then $0=0$.
3. **First description**: The column space is the plane containing all combinations of the pivot columns $(1,2,3)$ and $(3,8,7)$, since the pivots are in columns 1 and 3. **Second description**: The column space contains all vectors with $b_3+b_2-5b_1=0$. That makes $A\boldsymbol{x}=\boldsymbol{b}$ solvable, so $\boldsymbol{b}$ is in the column space. *All columns of A pass this test $b_3+b_2-5b_1=0$. This is the equation for the plane in the first description.*
4. The special solutions have free variables $x_2=1, x_4=0$ and then $x_2=0, x_4=1$:

Special solutions to $A\boldsymbol{x}=\boldsymbol{0}$
Back substitution in $U\boldsymbol{x}=\boldsymbol{0}$

$$s_1=\begin{bmatrix}-2\\1\\0\\0\end{bmatrix} \qquad s_2=\begin{bmatrix}-2\\0\\-1\\1\end{bmatrix}$$

The nullspace $N(A)$ in $\mathbf{R}^4$ contains all $\boldsymbol{x}_n=c_1s_1+c_2s_2=(-2c_1-2c_2, c_1, -c_2, c_2)$.

5. One particular solution $\boldsymbol{x}_p$ has free variables = zero. Back substitute in $U\boldsymbol{x}=\boldsymbol{c}$:

Particular solution to $A\boldsymbol{x}_p=(0,6,-6)$
This vector $\boldsymbol{b}$ satisfies $b_3+b_2-5b_1=0$

$$\boldsymbol{x}_p=\begin{bmatrix}-9\\0\\3\\0\end{bmatrix}$$

The complete solution to $A\boldsymbol{x}=(0,6,-6)$ is $\boldsymbol{x}=\boldsymbol{x}_p+$ all $\boldsymbol{x}_n$.

6. In the reduced form R, the third column changes from $(3,2,0)$ in U to $(0,1,0)$. The right side $\boldsymbol{c}=(0,6,0)$ now becomes $\boldsymbol{d}=(-9,3,0)$ showing -9 and 3 in $\boldsymbol{x}_p$:

$$[\,U \quad \boldsymbol{c}\,]=\begin{bmatrix}1&2&3&5&\mathbf{0}\\0&0&2&2&\mathbf{6}\\0&0&0&0&\mathbf{0}\end{bmatrix}\longrightarrow[\,R \quad \boldsymbol{d}\,]=\begin{bmatrix}1&2&0&2&\mathbf{-9}\\0&0&1&1&\mathbf{3}\\0&0&0&0&\mathbf{0}\end{bmatrix}$$

3.4 B If you have this information about the solutions to $A\boldsymbol{x}=\boldsymbol{b}$ for a specific $\boldsymbol{b}$, what does that tell you about the shape of A (and A itself)? And possibly about $\boldsymbol{b}$.

1. There is exactly one solution.
2. All solutions to $A\boldsymbol{x}=\boldsymbol{b}$ have the form $\boldsymbol{x}=\begin{bmatrix}1\\2\end{bmatrix}+c\begin{bmatrix}1\\1\end{bmatrix}$.
3. There are no solutions.
4. All solutions to $A\boldsymbol{x}=\boldsymbol{b}$ have the form $\boldsymbol{x}=\begin{bmatrix}1\\1\\0\end{bmatrix}+c\begin{bmatrix}1\\0\\1\end{bmatrix}$
5. There are infinitely many solutions.

Solution In case **1**, with exactly one solution, A must have full column rank $r=n$. The nullspace of A contains only the zero vector. Necessarily $m \geq n$.

In case **2**, A must have $n=2$ columns (and m is arbitrary). With $\begin{bmatrix}1\\1\end{bmatrix}$ in the nullspace of A, column 2 is the *negative* of column 1. With $x=\begin{bmatrix}1\\2\end{bmatrix}$ as a solution, $\boldsymbol{b}=(\text{column }1)+2\,(\text{column }2)=\text{column }2$. The columns can't be zero vectors.

In case **3** we only know that $\boldsymbol{b}$ is not in the column space of A. The rank of A must be less than m. I guess we know $\boldsymbol{b} \neq \boldsymbol{0}$, otherwise $\boldsymbol{x} = \boldsymbol{0}$ would be a solution.

In case **4**, A must have $n = 3$ columns. With $(1, 0, 1)$ in the nullspace of A, column 3 is the negative of column 1. Column 2 must *not* be a multiple of column 1, or the nullspace would contain another special solution. So the rank of A is $3 - 1 = 2$. Necessarily A has $m \geq 2$ rows. The right side $\boldsymbol{b}$ is column 1 + column 2.

In case **5** with infinitely many solutions, the nullspace must contain nonzero vectors. The rank r must be less than n (not full column rank), and $\boldsymbol{b}$ must be in the column space of A. We don't know if *every* $\boldsymbol{b}$ is in the column space, so we don't know if $r = m$.

3.4 C Find the complete solution $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n$ by forward elimination on $[A \ \ \boldsymbol{b}]$:

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 4 & 8 \\ 4 & 8 & 6 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 10 \end{bmatrix}.$$

Find numbers y_1, y_2, y_3 so that y_1 (row 1)$+y_2$ (row 2)$+y_3$ (row 3) $=$ ***zero row***. Check that $\boldsymbol{b} = (4, 2, 10)$ satisfies the condition $y_1b_1 + y_2b_2 + y_3b_3 = 0$. Why is this the condition for the equations to be solvable and $\boldsymbol{b}$ to be in the column space?

Solution Forward elimination on $[A \ \ \boldsymbol{b}]$ produces a zero row in $[U \ \ \boldsymbol{c}]$. The third equation becomes $0 = 0$ and the equations are consistent (and solvable):

$$\begin{bmatrix} 1 & 2 & 1 & 0 & \mathbf{4} \\ 2 & 4 & 4 & 8 & \mathbf{2} \\ 4 & 8 & 6 & 8 & \mathbf{10} \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 & 0 & \mathbf{4} \\ 0 & 0 & 2 & 8 & \mathbf{-6} \\ 0 & 0 & 2 & 8 & \mathbf{-6} \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 & 0 & \mathbf{4} \\ 0 & 0 & 2 & 8 & \mathbf{-6} \\ 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}.$$

Columns 1 and 3 contain pivots. The variables x_2 and x_4 are free. If we set those to zero we can solve (back substitution) for the particular solution $\boldsymbol{x}_p = (7, 0, -3, 0)$. We see 7 and -3 again if elimination continues all the way to $[R \ \ \boldsymbol{d}]$:

$$\begin{bmatrix} 1 & 2 & 1 & 0 & \mathbf{4} \\ 0 & 0 & 2 & 8 & \mathbf{-6} \\ 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 & 0 & \mathbf{4} \\ 0 & 0 & 1 & 4 & \mathbf{-3} \\ 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 0 & -4 & \mathbf{7} \\ 0 & 0 & 1 & 4 & \mathbf{-3} \\ 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}.$$

For the nullspace part $\boldsymbol{x}_n$ with $\boldsymbol{b} = \boldsymbol{0}$, set the free variables x_2, x_4 to $1, 0$ and also $0, 1$:

Special solutions $\quad s_1 = (-2, 1, 0, 0)$ **and** $s_2 = (4, 0, -4, 1)$

Then the complete solution to $A\boldsymbol{x} = \boldsymbol{b}$ (and $R\boldsymbol{x} = \boldsymbol{d}$) is $\boldsymbol{x}_{\text{complete}} = \boldsymbol{x}_p + c_1 s_1 + c_2 s_2$.

The rows of A produced the zero row from 2(row 1) + (row 2) $-$ (row 3) $=$ $(0, 0, 0, 0)$. The same combination for $\boldsymbol{b} = (4, 2, 10)$ gives $2(4) + (2) - (10) = 0$. If a combination of the rows (on the left side) gives the zero row, then the same combination must give zero on the right side. Of course! Otherwise no solution.

Later we will say this again in different words: If every column of A is perpendicular to $\boldsymbol{y} = (2, 1, -1)$, then any combination $\boldsymbol{b}$ of those columns must also be perpendicular to $\boldsymbol{y}$. Otherwise $\boldsymbol{b}$ is not in the column space and $A\boldsymbol{x} = \boldsymbol{b}$ is not solvable.

And again: If $\boldsymbol{y}$ is in the nullspace of A^{T} then $\boldsymbol{y}$ must be perpendicular to every $\boldsymbol{b}$ in the column space. Just looking ahead ...

Problem Set 3.4

1 (Recommended) Execute the six steps of Worked Example **3.4 A** to describe the column space and nullspace of A and the complete solution to $A\boldsymbol{x} = \boldsymbol{b}$:

$$A = \begin{bmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}$$

2 Carry out the same six steps for this matrix A with rank one. You will find *two* conditions on b_1, b_2, b_3 for $A\boldsymbol{x} = \boldsymbol{b}$ to be solvable. Together these two conditions put $\boldsymbol{b}$ into the _____ space (two planes give a line):

$$A = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} [2 \;\; 1 \;\; 3] = \begin{bmatrix} 2 & 1 & 3 \\ 6 & 3 & 9 \\ 4 & 2 & 6 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 30 \\ 20 \end{bmatrix}$$

Questions 3–15 are about the solution of $A\boldsymbol{x} = \boldsymbol{b}$. Follow the steps in the text to $\boldsymbol{x}_p$ and $\boldsymbol{x}_n$. Use the augmented matrix with last column $\boldsymbol{b}$.

3 Write the complete solution as $\boldsymbol{x}_p$ plus any multiple of $\boldsymbol{s}$ in the nullspace:

$$\begin{aligned} x + 3y + 3z &= 1 \\ 2x + 6y + 9z &= 5 \\ -x - 3y + 3z &= 5. \end{aligned}$$

4 Find the complete solution (also called the *general solution*) to

$$\begin{bmatrix} 1 & 3 & 1 & 2 \\ 2 & 6 & 4 & 8 \\ 0 & 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}.$$

5 Under what condition on b_1, b_2, b_3 is this system solvable? Include $\boldsymbol{b}$ as a fourth column in elimination. Find all solutions when that condition holds:

$$\begin{aligned} x + 2y - 2z &= b_1 \\ 2x + 5y - 4z &= b_2 \\ 4x + 9y - 8z &= b_3. \end{aligned}$$

6 What conditions on b_1, b_2, b_3, b_4 make each system solvable? Find $\boldsymbol{x}$ in that case:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 2 & 5 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 5 & 7 \\ 3 & 9 & 12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}.$$

7 Show by elimination that (b_1, b_2, b_3) is in the column space if $b_3 - 2b_2 + 4b_1 = 0$.

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 8 & 2 \\ 2 & 4 & 0 \end{bmatrix}.$$

What combination of the rows of A gives the zero row?

8 Which vectors (b_1, b_2, b_3) are in the column space of A? Which combinations of the rows of A give zero?

(a) $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & 3 \\ 0 & 2 & 5 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix}$.

9 (a) The Worked Example **3.4 A** reached $[\,U \;\; \boldsymbol{c}\,]$ from $[\,A \;\; \boldsymbol{b}\,]$. Put the multipliers into L and verify that LU equals A and $L\boldsymbol{c}$ equals $\boldsymbol{b}$.

(b) Combine the pivot columns of A with the numbers -9 and 3 in the particular solution $\boldsymbol{x}_p$. What is that linear combination and why?

10 Construct a 2 by 3 system $A\boldsymbol{x} = \boldsymbol{b}$ with particular solution $\boldsymbol{x}_p = (2, 4, 0)$ and homogeneous solution $\boldsymbol{x}_n =$ any multiple of $(1, 1, 1)$.

11 Why can't a 1 by 3 system have $\boldsymbol{x}_p = (2, 4, 0)$ and $\boldsymbol{x}_n =$ any multiple of $(1, 1, 1)$?

12 (a) If $A\boldsymbol{x} = \boldsymbol{b}$ has two solutions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, find two solutions to $A\boldsymbol{x} = \boldsymbol{0}$.

(b) Then find another solution to $A\boldsymbol{x} = \boldsymbol{0}$ and another solution to $A\boldsymbol{x} = \boldsymbol{b}$.

13 Explain why these are all false:

(a) The complete solution is any linear combination of $\boldsymbol{x}_p$ and $\boldsymbol{x}_n$.

(b) A system $A\boldsymbol{x} = \boldsymbol{b}$ has at most one particular solution.

(c) The solution $\boldsymbol{x}_p$ with all free variables zero is the shortest solution (minimum length $\|\boldsymbol{x}\|$). Find a 2 by 2 counterexample.

(d) If A is invertible there is no solution $\boldsymbol{x}_n$ in the nullspace.

14 Suppose column 5 of U has no pivot. Then x_5 is a _____ variable. The zero vector (is) (is not) the only solution to $A\boldsymbol{x} = \boldsymbol{0}$. If $A\boldsymbol{x} = \boldsymbol{b}$ has a solution, then it has _____ solutions.

15 Suppose row 3 of U has no pivot. Then that row is _____. The equation $U\boldsymbol{x} = \boldsymbol{c}$ is only solvable provided _____. The equation $A\boldsymbol{x} = \boldsymbol{b}$ (*is*) (*is not*) (*might not be*) solvable.

Questions 16–20 are about matrices of "full rank" $r = m$ or $r = n$.

16 The largest possible rank of a 3 by 5 matrix is _____. Then there is a pivot in every _____ of U and R. The solution to $A\boldsymbol{x} = \boldsymbol{b}$ (*always exists*) (*is unique*). The column space of A is _____. An example is $A =$ _____.

17 The largest possible rank of a 6 by 4 matrix is _____. Then there is a pivot in every _____ of U and R. The solution to $A\boldsymbol{x} = \boldsymbol{b}$ (*always exists*) (*is unique*). The nullspace of A is _____. An example is $A =$ _____.

18 Find by elimination the rank of A and also the rank of A^{T}:

$$A = \begin{bmatrix} 1 & 4 & 0 \\ 2 & 11 & 5 \\ -1 & 2 & 10 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & q \end{bmatrix} \quad \text{(rank depends on } q\text{)}.$$

19 Find the rank of A and also of $A^{\mathrm{T}}A$ and also of AA^{T}:

$$A = \begin{bmatrix} 1 & 1 & 5 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

20 Reduce A to its echelon form U. Then find a triangular L so that $A = LU$.

$$A = \begin{bmatrix} 3 & 4 & 1 & 0 \\ 6 & 5 & 2 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & 2 & 0 & 3 \\ 0 & 6 & 5 & 4 \end{bmatrix}.$$

21 Find the complete solution in the form $\boldsymbol{x}_p + \boldsymbol{x}_n$ to these full rank systems:

(a) $x + y + z = 4$ (b) $\begin{aligned} x + y + z &= 4 \\ x - y + z &= 4. \end{aligned}$

22 If $A\boldsymbol{x} = \boldsymbol{b}$ has infinitely many solutions, why is it impossible for $A\boldsymbol{x} = \boldsymbol{B}$ (new right side) to have only one solution? Could $A\boldsymbol{x} = \boldsymbol{B}$ have no solution?

23 Choose the number q so that (if possible) the ranks are (a) 1, (b) 2, (c) 3:

$$A = \begin{bmatrix} 6 & 4 & 2 \\ -3 & -2 & -1 \\ 9 & 6 & q \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 1 & 3 \\ q & 2 & q \end{bmatrix}.$$

24 Give examples of matrices A for which the number of solutions to $A\boldsymbol{x} = \boldsymbol{b}$ is

(a) 0 or 1, depending on $\boldsymbol{b}$

(b) ∞, regardless of $\boldsymbol{b}$

(c) 0 or ∞, depending on $\boldsymbol{b}$

(d) 1, regardless of $\boldsymbol{b}$.

25 Write down all known relations between r and m and n if $A\boldsymbol{x} = \boldsymbol{b}$ has

(a) no solution for some $\boldsymbol{b}$

(b) infinitely many solutions for every $\boldsymbol{b}$

(c) exactly one solution for some $\boldsymbol{b}$, no solution for other $\boldsymbol{b}$

(d) exactly one solution for every $\boldsymbol{b}$.

Questions 26–33 are about Gauss-Jordan elimination (upwards as well as downwards) and the reduced echelon matrix R.

26 Continue elimination from U to R. Divide rows by pivots so the new pivots are all 1. Then produce zeros *above* those pivots to reach R:

$$U = \begin{bmatrix} 2 & 4 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 2 & 4 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix}.$$

27 Suppose U is square with n pivots (an invertible matrix). *Explain why* $R = I$.

28 Apply Gauss-Jordan elimination to $U\boldsymbol{x} = \boldsymbol{0}$ and $U\boldsymbol{x} = \boldsymbol{c}$. Reach $R\boldsymbol{x} = \boldsymbol{0}$ and $R\boldsymbol{x} = \boldsymbol{d}$:

$$\begin{bmatrix} U & \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & \mathbf{0} \\ 0 & 0 & 4 & \mathbf{0} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} U & \boldsymbol{c} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & \mathbf{5} \\ 0 & 0 & 4 & \mathbf{8} \end{bmatrix}.$$

Solve $R\boldsymbol{x} = \boldsymbol{0}$ to find $\boldsymbol{x}_n$ (its free variable is $x_2 = 1$). Solve $R\boldsymbol{x} = \boldsymbol{d}$ to find $\boldsymbol{x}_p$ (its free variable is $x_2 = 0$).

29 Apply Gauss-Jordan elimination to reduce to $R\boldsymbol{x} = \boldsymbol{0}$ and $R\boldsymbol{x} = \boldsymbol{d}$:

$$\begin{bmatrix} U & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 6 & \mathbf{0} \\ 0 & 0 & 2 & \mathbf{0} \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} U & \boldsymbol{c} \end{bmatrix} = \begin{bmatrix} 3 & 0 & 6 & \mathbf{9} \\ 0 & 0 & 2 & \mathbf{4} \\ 0 & 0 & 0 & \mathbf{5} \end{bmatrix}.$$

Solve $U\boldsymbol{x} = \boldsymbol{0}$ or $R\boldsymbol{x} = \boldsymbol{0}$ to find $\boldsymbol{x}_n$ (free variable $= 1$). What are the solutions to $R\boldsymbol{x} = \boldsymbol{d}$?

30 Reduce to $Ux = c$ (Gaussian elimination) and then $Rx = d$ (Gauss-Jordan):

$$Ax = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 1 & 3 & 2 & 0 \\ 2 & 0 & 4 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 10 \end{bmatrix} = b.$$

Find a particular solution x_p and all homogeneous solutions x_n.

31 Find matrices A and B with the given property or explain why you can't: The only solution of $Ax = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is $x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The only solution of $Bx = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

32 Find the LU factorization of A and the complete solution to $Ax = b$:

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 5 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 5 \end{bmatrix} \quad \text{and then} \quad b = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

33 The complete solution to $Ax = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Find A.

34 Suppose you know that the 3 by 4 matrix A has the vector $s = (2, 3, 1, 0)$ as a basis for its nullspace

(a) What is the *rank* of A and the complete solution to $Ax = \mathbf{0}$?

(b) What is the exact row reduced echelon form R of A?

SINGULAR VALUE DECOMPOSITION (SVD) ■ 6.7

The Singular Value Decomposition is a highlight of linear algebra. A is any m by n matrix, square or rectangular. We will diagonalize it, but not by $S^{-1}AS$. Its row space is r-dimensional (inside $\mathbf{R}^n$). Its column space is also r-dimensional (inside $\mathbf{R}^m$). We are going to choose special ***orthonormal bases*** $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ for the row space and $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ for the column space. For those bases, we want each $A\boldsymbol{v}_i$ to be in the direction of $\boldsymbol{u}_i$. In matrix form, ***these equations*** $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$ ***become*** $AV = U\Sigma$ ***or*** $A = U\Sigma V^{\mathrm{T}}$. This is the SVD.

Image Compression

Unusually, I am going to stop the theory and describe applications. This is the century of data, and often that data is stored in a matrix. *A digital image is really a matrix of pixel values.* Each little picture element or "pixel" has a gray scale number between black and white (it has three numbers for a color picture). The picture might have $512 = 2^9$ pixels in each row and $256 = 2^8$ pixels down each column. We have a 256 by 512 pixel matrix with 2^{17} entries! To store one picture, the computer has no problem. But if you go in for a CT scan or Magnetic Resonance, you produce an image at every cross section—a ton of data. If the pictures are frames in a movie, 30 frames a second means 108,000 images per hour. Compression is especially needed for high definition digital TV, or the equipment could not keep up in real time.

What is compression? We want to replace those 2^{17} matrix entries by a smaller number, *without losing picture quality.* A simple way would be to use larger pixels—replace groups of four pixels by their average value. This is 4 : 1 compression. But if we carry it further, like 16 : 1, our image becomes "blocky". We want to replace the mn entries by a smaller number, in a way that the human visual system won't notice.

Compression is a billion dollar problem and everyone has ideas. Later in this book I will describe Fourier transforms (used in **jpeg**) and wavelets (now in **JPEG2000**). Here we try an SVD approach: *Replace the* 256 *by* 512 *pixel matrix by* ***a matrix of rank one****: a column times a row.* If this is successful, the storage requirement for a column and row becomes $256 + 512$ (plus instead of times!). The compression ratio $(256)(512)/(256 + 512)$ is better than 170 : 1. This is more than we hope for. We may actually use five matrices of rank one (so a matrix approximation of rank 5). The compression is still 34 : 1 and the crucial question is the picture quality.

Where does the SVD come in? *The best rank one approximation to A is the matrix $\sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}}$.* It uses the largest singular value σ_1 and its left and right singular vectors $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$. The best rank 5 approximation includes $\sigma_2 \boldsymbol{u}_2 \boldsymbol{v}_2^{\mathrm{T}} + \cdots + \sigma_5 \boldsymbol{u}_5 \boldsymbol{v}_5^{\mathrm{T}}$. If we can compute those $\boldsymbol{u}$'s and $\boldsymbol{v}$'s quickly (a big "if" since you will see them as eigenvectors for $A^{\mathrm{T}}A$ and AA^{T}) then this SVD algorithm is competitive.

I will mention a different matrix, one that a library needs to compress. The rows correspond to key words. The columns correspond to titles in the library. The entry in this *word-title matrix* is $a_{ij} = 1$ if word i is in title j (otherwise $a_{ij} = 0$). We might normalize the columns to be unit vectors, so that long titles don't get an advantage.

Instead of the title, we might use a table of contents or an abstract that better captures the content. (Other books might share the title "*Introduction to Linear Algebra*". If you are searching for the SVD, you want the right book.) Instead of $a_{ij} = 1$, the entries of A can include the *frequency* of the search words in each document.

Once the indexing matrix is created, the search is a linear algebra problem. If we use 100,000 words from an English dictionary and 2,000,000,000 web pages as documents, it is a long search. We need a shortcut. This matrix has to be compressed. I will now explain the SVD approach, which gives an optimal low rank approximation to A. (It works better for library matrices than for natural images.) There is an ever-present tradeoff in the cost to compute the u's and v's, and I hope you will invent a better way.

The Bases and the SVD

Start with a 2 by 2 matrix. Let its rank be $r = 2$, so this matrix A is invertible. Its row space is the plane $\mathbf{R}^2$. We want $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ to be perpendicular unit vectors, an orthonormal basis. ***We also want $A\boldsymbol{v}_1$ and $A\boldsymbol{v}_2$ to be perpendicular.*** (This is the tricky part. It is what makes the bases special.) Then the unit vectors $\boldsymbol{u}_1 = A\boldsymbol{v}_1/\|A\boldsymbol{v}_1\|$ and $\boldsymbol{u}_2 = A\boldsymbol{v}_2/\|A\boldsymbol{v}_2\|$ will be orthonormal. As a specific example, we work with the unsymmetric matrix

$$A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}. \tag{1}$$

First point Why not choose one orthogonal basis in Q, instead of two in U and V? *Because no orthogonal matrix Q will make $Q^{-1}AQ$ diagonal.* We need $U^{-1}AV$.

Second point Why not choose the eigenvectors of A as the basis? *Because that basis is not orthonormal.* A is not symmetric and we need *two different* orthogonal matrices.

We are aiming for orthonormal bases that diagonalize A. The two bases will be different—one basis cannot do it. When the inputs are $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, the outputs are $A\boldsymbol{v}_1$ and $A\boldsymbol{v}_2$. We want those to line up with $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. ***The basis vectors have to give $A\boldsymbol{v}_1 = \sigma_1\boldsymbol{u}_1$ and also $A\boldsymbol{v}_2 = \sigma_2\boldsymbol{u}_2$.*** The "singular values" σ_1 and σ_2 are the lengths $\|A\boldsymbol{v}_1\|$ and $\|A\boldsymbol{v}_2\|$. With $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ as columns of V you see what we are asking for:

$$A\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} \sigma_1\boldsymbol{u}_1 & \sigma_2\boldsymbol{u}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix}\begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix}. \tag{2}$$

In matrix notation that is $AV = U\Sigma$, or $U^{-1}AV = \Sigma$, or $U^{\mathrm{T}}AV = \Sigma$. The diagonal matrix Σ is like Λ (capital sigma versus capital lambda). Σ contains the *singular values* σ_1, σ_2, which are different from the eigenvalues λ_1, λ_2 in Λ.

The difference comes from U and V. When they both equal S, we have $S^{-1}AS = \Lambda$. The matrix is diagonalized. But the eigenvectors in S are not generally orthonormal. The new requirement is that *U and V must be orthogonal matrices.*

$$\textbf{Orthonormal basis} \qquad V^{\mathrm{T}}V = \begin{bmatrix} - & v_1^{\mathrm{T}} & - \\ - & v_2^{\mathrm{T}} & - \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3}$$

Thus $V^{\mathrm{T}}V = I$ which means $V^{\mathrm{T}} = V^{-1}$. Similarly $U^{\mathrm{T}}U = I$ and $U^{\mathrm{T}} = U^{-1}$.

6R The ***Singular Value Decomposition*** **(SVD)** has orthogonal matrices U and V:

$$AV = U\Sigma \quad \text{and then} \quad A = U\Sigma V^{-1} = U\Sigma V^{\mathrm{T}}. \tag{4}$$

This is the new factorization of A: ***orthogonal*** times ***diagonal*** times ***orthogonal***.

There is a neat way to remove U and see V by itself: *Multiply A^{T} times A.*

$$A^{\mathrm{T}}A = (U\Sigma V^{\mathrm{T}})^{\mathrm{T}}(U\Sigma V^{\mathrm{T}}) = V\Sigma^{\mathrm{T}}\Sigma V^{\mathrm{T}}. \tag{5}$$

$U^{\mathrm{T}}U$ disappears because it equals I. Then Σ^{T} is next to Σ. Multiplying those diagonal matrices gives σ_1^2 and σ_2^2. That leaves an ordinary diagonalization of the crucial symmetric matrix $A^{\mathrm{T}}A$, whose eigenvalues are σ_1^2 and σ_2^2:

$$A^{\mathrm{T}}A = V \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} V^{\mathrm{T}}. \tag{6}$$

This is exactly like $A = Q\Lambda Q^{\mathrm{T}}$. But the symmetric matrix is not A itself. Now the symmetric matrix is $A^{\mathrm{T}}A$! *And the columns of V are the eigenvectors of $A^{\mathrm{T}}A$.*

This tells us how to find V. We are ready to complete the example.

Example 1 Find the singular value decomposition of $A = \left[\begin{smallmatrix} 2 & 2 \\ -1 & 1 \end{smallmatrix}\right]$.

Solution Compute $A^{\mathrm{T}}A$ and its eigenvectors. Then make them unit vectors:

$$A^{\mathrm{T}}A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \quad \text{has unit eigenvectors} \quad v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \text{and} \quad v_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

The eigenvalues of $A^{\mathrm{T}}A$ are 8 and 2. The v's are perpendicular, because eigenvectors of every symmetric matrix are perpendicular—and $A^{\mathrm{T}}A$ is automatically symmetric.

What about u_1 and u_2? They are quick to find, because Av_1 is going to be in the direction of u_1 and Av_2 is in the direction of u_2:

$$Av_1 = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}. \quad \text{The unit vector is} \quad u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

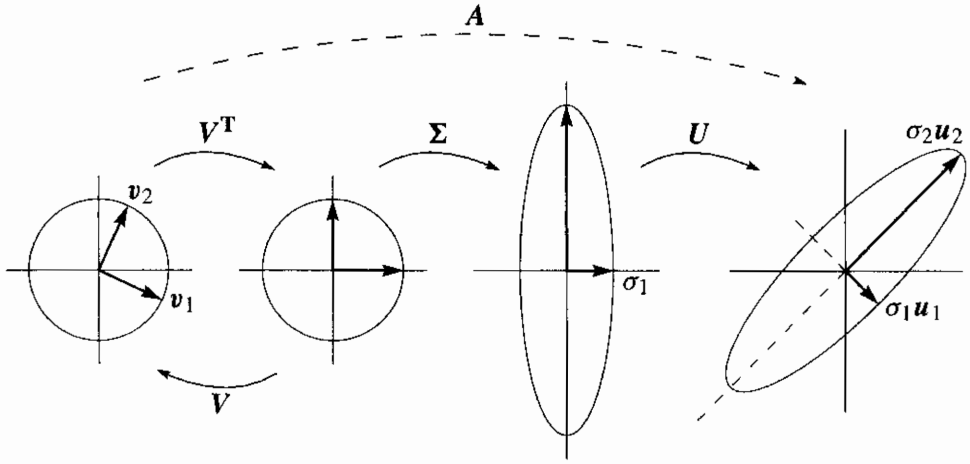


Figure 6.5 U and V are rotations and reflections. Σ is a stretching matrix.

Clearly $A\boldsymbol{v}_1$ is the same as $2\sqrt{2}\boldsymbol{u}_1$. The first singular value is $\sigma_1 = 2\sqrt{2}$. Then $\sigma_1^2 = 8$, which is the eigenvalue of $A^{\mathrm{T}}A$. We have $A\boldsymbol{v}_1 = \sigma_1\boldsymbol{u}_1$ exactly as required. Similarly

$$A\boldsymbol{v}_2 = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}. \quad \text{The unit vector is} \quad \boldsymbol{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Now $A\boldsymbol{v}_2$ is $\sqrt{2}\boldsymbol{u}_2$. The second singular value is $\sigma_2 = \sqrt{2}$. And σ_2^2 agrees with the other eigenvalue 2 of $A^{\mathrm{T}}A$. We have completed the SVD:

$$A = U\Sigma V^{\mathrm{T}} \quad \text{is} \quad \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2\sqrt{2} & \\ & \sqrt{2} \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}. \tag{7}$$

This matrix, and every invertible 2 by 2 matrix, ***transforms the unit circle to an ellipse***. You can see that in the figure, which was created by Cliff Long and Tom Hern.

One final point about that example. We found the $\boldsymbol{u}$'s from the $\boldsymbol{v}$'s. Could we find the $\boldsymbol{u}$'s directly? *Yes*, by multiplying AA^{T} instead of $A^{\mathrm{T}}A$:

$$AA^{\mathrm{T}} = (U\Sigma V^{\mathrm{T}})(V\Sigma^{\mathrm{T}}U^{\mathrm{T}}) = U\Sigma\Sigma^{\mathrm{T}}U^{\mathrm{T}}. \tag{8}$$

This time it is $V^{\mathrm{T}}V = I$ that disappears. Multiplying $\Sigma\Sigma^{\mathrm{T}}$ gives σ_1^2 and σ_2^2 as before. *The columns of U are the eigenvectors of AA^{T}*:

$$AA^{\mathrm{T}} = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}.$$

This matrix happens to be diagonal. Its eigenvectors are $(1, 0)$ and $(0, 1)$. This agrees with $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ found earlier. Why should we take the first eigenvector to be $(1, 0)$ instead of $(0, 1)$? Because we have to follow the order of the eigenvalues. Notice that AA^{T} has the same eigenvalues (8 and 2) as $A^{\mathrm{T}}A$. The singular values are $\sqrt{8}$ and $\sqrt{2}$.

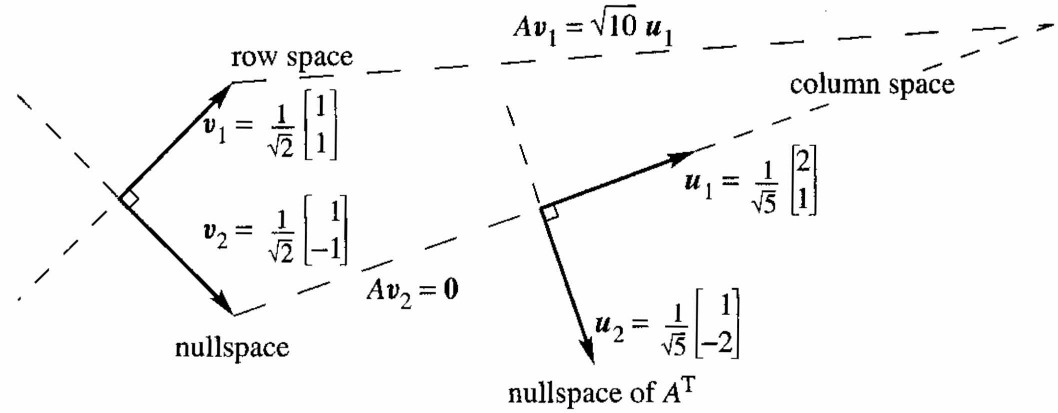


Figure 6.6 The SVD chooses orthonormal bases for 4 subspaces so that $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$.

Example 2 Find the SVD of the singular matrix $A = \begin{bmatrix}2&2\\1&1\end{bmatrix}$. The rank is $r = 1$. The row space has only one basis vector $\boldsymbol{v}_1$. The column space has only one basis vector $\boldsymbol{u}_1$. We can see those vectors $(1, 1)$ and $(2, 1)$ in A, and make them into unit vectors:

$$\textbf{Row space} \quad \boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} \qquad\qquad \textbf{Column space} \quad \boldsymbol{u}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix}2\\1\end{bmatrix}$$

Then $A\boldsymbol{v}_1$ must equal $\sigma_1\boldsymbol{u}_1$. It does, with $\sigma_1 = \sqrt{10}$. This A is $\sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^T$ with rank 1.

The SVD could stop after the row basis and column basis (it usually doesn't). It is customary for U and V to be square. The matrices need a second column. The vector $\boldsymbol{v}_2$ must be orthogonal to $\boldsymbol{v}_1$, and $\boldsymbol{u}_2$ must be orthogonal to $\boldsymbol{u}_1$:

$$\boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \qquad \text{and} \qquad \boldsymbol{u}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\-2\end{bmatrix}.$$

The vector $\boldsymbol{v}_2$ is in the nullspace. It is perpendicular to $\boldsymbol{v}_1$ in the row space. Multiply by A to get $A\boldsymbol{v}_2 = \boldsymbol{0}$. We could say that the second singular value is $\sigma_2 = 0$, but singular values are like pivots—only the r nonzeros are counted.

All three matrices U, Σ, V are 2 by 2 in the complete SVD:

$$\begin{bmatrix}2&2\\1&1\end{bmatrix} = U\Sigma V^T = \frac{1}{\sqrt{5}}\begin{bmatrix}2&1\\1&-2\end{bmatrix}\begin{bmatrix}\sqrt{10}&0\\0&0\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}. \tag{9}$$

6S ***The matrices U and V contain orthonormal bases for all four subspaces:***

first	r	columns of V :	*row space* of A
last	$n-r$	columns of V :	*nullspace* of A
first	r	columns of U :	*column space* of A
last	$m-r$	columns of U :	*nullspace* of A^T.

The first columns $v_1, \ldots, v_r$ and $u_1, \ldots, u_r$ are eigenvectors of A^TA and AA^T. Then Av_i falls in the direction of u_i, and we now explain why. The last v's and u's (in the nullspaces) are easier. As long as those are orthonormal, the SVD will be correct.

Proof of SVD: Start from $A^TAv_i = \sigma_i^2 v_i$, which gives the v's and σ's. To prove that $Av_i = \sigma_i u_i$, the key steps are to multiply by v_i^T and by A:

$$v_i^T A^T A v_i = \sigma_i^2 v_i^T v_i \quad \text{gives} \quad \|Av_i\|^2 = \sigma_i^2 \quad \text{so that} \quad \|Av_i\| = \sigma_i \tag{10}$$

$$AA^T A v_i = \sigma_i^2 A v_i \quad \text{gives} \quad u_i = Av_i/\sigma_i \quad \text{as a unit eigenvector of} \quad AA^T. \tag{11}$$

Equation (10) used the small trick of placing parentheses in $(v_i^T A^T)(Av_i)$. This is a vector Av_i times its transpose, giving $\|Av_i\|^2$. Equation (11) placed the parentheses in $(AA^T)(Av_i)$. This shows that Av_i is an eigenvector of AA^T. We divide by its length σ_i to get the unit vector $u_i = Av_i/\sigma_i$. This is the equation $Av_i = \sigma_i u_i$ that we want! It says that A is diagonalized by these outstanding bases.

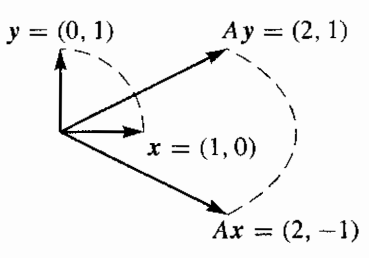
I will give you my opinion directly. The SVD is the climax of this linear algebra course. I think of it as the final step in the Fundamental Theorem. First come the *dimensions* of the four subspaces. Then their *orthogonality*. Then the *orthonormal bases which diagonalize A*. It is all in the formula $A = U\Sigma V^T$. More applications are coming—they are certainly important—but you have made it to the top.

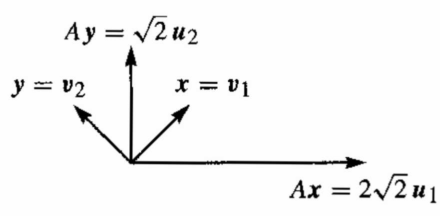
Eigshow (Part 2)

Section 6.1 described the MATLAB demo called **eigshow**. The first option is *eig*, when x moves in a circle and Ax follows on an ellipse. The second option is *svd*, when two vectors x and y stay perpendicular as they travel around a circle. Then Ax and Ay move too (not usually perpendicular). There are four vectors on the screen.

The SVD is seen graphically when Ax is perpendicular to Ay. Their directions at that moment give an orthonormal basis u_1, u_2. Their lengths give the singular values σ_1, σ_2. The vectors x and y at that same moment are the orthonormal basis v_1, v_2.

The Java demo on **web.mit.edu/18.06/www** shows $Av_1 = \sigma_1 u_1$ and $Av_2 = \sigma_2 u_2$. In matrix language that is $AV = U\Sigma$. This is the SVD.





Searching the Web

I will end with an application of the SVD to web search engines. When you type a search word, you get a list of related web sites in order of importance. (Regrettably, typing "SVD" produced 13 non-mathematical SVD's before the real one. "Cofactors" was even worse but "cofactor" had one good entry. "Four subspaces" did much better.) The HITS algorithm that we describe is one way to produce that ranked list. It begins with about 200 sites found from an index of key words, and after that we look only at links between pages. HITS is link-based not content-based.

Start with the 200 sites and all sites that link to them and all sites they link to. That is our list, to be put in order. Importance can be measured in two ways:

1. The site is an ***authority***: *links come from many sites.* Especially from hubs.

2. The site is a ***hub***: *it links to many sites in the list.* Especially to authorities.

We want numbers $x_1, \dots, x_N$ to rank the authorities and $y_1, \dots, y_N$ to rank the hubs. Start with a simple count: x_i^0 and y_i^0 count the links into and out of site i.

Here is the point: *A good authority has links from important sites* (like hubs). Links from universities count more heavily than links from friends. *A good hub is linked to important sites* (like authorities). A link to amazon.com means more than a link to wellesleycambridge.com. The rankings x^0 and y^0 from counting links are updated to x^1 and y^1 by taking account of *good* links (measuring their quality by x^0 and y^0):

$$\textbf{Authority values} \quad x_i^1 = \sum_{j \text{ links to } i} y_j^0 \qquad\qquad \textbf{Hub values} \quad y_i^1 = \sum_{i \text{ links to } j} x_j^0 \tag{12}$$

In matrix language those are $x^1 = A^{\mathrm{T}}y^0$ and $y^1 = Ax^0$. The matrix A contains 1's and 0's, with $a_{ij} = 1$ when i links to j. In the language of graphs, A is an "adjacency matrix" for the World Wide Web. It is pretty large.

The algorithm doesn't stop there. The new x^1 and y^1 give better rankings, but not the best. Take another step like (12) to x^2 and y^2. *Notice how $A^{\mathrm{T}}A$ and AA^{T} appear*:

$$x^2 = A^{\mathrm{T}}y^1 = A^{\mathrm{T}}Ax^0 \qquad \text{and} \qquad y^2 = A^{\mathrm{T}}x^1 = AA^{\mathrm{T}}y^0. \tag{13}$$

In two steps we are multiplying x^0 by $A^{\mathrm{T}}A$ and y^0 by AA^{T}. In twenty steps we are multiplying by $(A^{\mathrm{T}}A)^{10}$ and $(AA^{\mathrm{T}})^{10}$. When we take these powers, the largest eigenvalue σ_1^2 begins to dominate. And the vectors x and y gradually line up with the leading eigenvectors v_1 and u_1 of $A^{\mathrm{T}}A$ and AA^{T}. We are computing the top terms in the SVD iteratively, by the *power method* that is further discussed in Section 9.3. It is wonderful that linear algebra helps to understand the Web.

Google actually creates rankings by a random walk that follows web links. The more often this random walk goes to a site, the higher the ranking. The frequency of visits gives the leading eigenvector ($\lambda = 1$) of the normalized adjacency matrix for the

Web. ***That matrix has 2.7 billion rows and columns, from 2.7 billion web sites.*** This is the largest eigenvalue problem ever solved.

Some details are on the Web, but many important techniques are secrets of *Google*: `www.mathworks.com/company/newsletter/clevescorner/oct02_cleve.shtml` Probably *Google* starts with last month's eigenvector as a first approximation, and runs the random walk very fast. To get a high ranking, you want a lot of links from important sites. The HITS algorithm is described in the 1999 *Scientific American* (June 16). But I don't think the SVD is mentioned there. . .

■ REVIEW OF THE KEY IDEAS ■

1. The SVD factors A into $U\Sigma V^{\mathrm{T}}$, with r singular values $\sigma_1 \geq \ldots \geq \sigma_r > 0$.

2. The numbers $\sigma_1^2, \ldots, \sigma_r^2$ are the nonzero eigenvalues of AA^{T} and $A^{\mathrm{T}}A$.

3. The orthonormal columns of U and V are eigenvectors of AA^{T} and $A^{\mathrm{T}}A$.

4. Those columns are orthonormal bases for the four fundamental subspaces of A.

5. Those bases diagonalize the matrix: $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$ for $i \leq r$. This is $AV = U\Sigma$.

■ WORKED EXAMPLES ■

6.7 A Identify by name these decompositions $A = \boldsymbol{c}_1\boldsymbol{r}_1 + \cdots + \boldsymbol{c}_n\boldsymbol{r}_n$ of an n by n matrix into n rank one matrices (column $\boldsymbol{c}$ times row $\boldsymbol{r}$):

1. *Orthogonal* columns $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ and *orthogonal* rows $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$
2. *Orthogonal* columns $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ and *triangular* rows $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$
3. *Triangular* columns $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ and *triangular* rows $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$

Triangular means that $\boldsymbol{c}_i$ and $\boldsymbol{r}_i$ have zeros before component i. The matrix C with columns $\boldsymbol{c}_i$ is lower triangular, the matrix R with rows $\boldsymbol{r}_i$ is upper triangular. Where do the rank and the pivots and singular values come into this picture?

Solution These three splittings $A = CR$ are basic to linear algebra, pure or applied:

1. Singular Value Decomposition $A = U\Sigma V^{\mathrm{T}}$ (*orthogonal* U, *orthogonal* ΣV^{T})

2. Gram-Schmidt Orthogonalization $A = QR$ (*orthogonal* Q, *triangular* R)

3. Gaussian Elimination $A = LU$ (*triangular* L, *triangular* U)

When A (possibly rectangular) has rank r, we need only r rank one matrices (not n).

With orthonormal rows in V^{T}, the σ's in Σ come in: $A = \sigma_1 \boldsymbol{c}_1 \boldsymbol{r}_1 + \cdots + \sigma_n \boldsymbol{c}_n \boldsymbol{r}_n$. With diagonal 1's in L and U, the pivots d_i come in: $A = LDU = d_1 \boldsymbol{c}_1 \boldsymbol{r}_1 + \cdots + d_n \boldsymbol{c}_n \boldsymbol{r}_n$. With the diagonal of R placed in H, QR becomes $QHR = h_1 \boldsymbol{c}_1 \boldsymbol{r}_1 + \cdots + h_n \boldsymbol{c}_n \boldsymbol{r}_n$. These numbers h_i have no standard name and I propose "heights". Each h_i tells the height of column i above the base from the first $i-1$ columns. The volume of the full n-dimensional box comes from $A = U\Sigma V^{\mathrm{T}} = LDU = QHR$:

$$|\det A| = |\,\textit{product of } \sigma\text{'s}\,| = |\,\textit{product of } d\text{'s}\,| = |\,\textit{product of } h\text{'s}\,|.$$

Problem Set 6.7

Problems 1–3 compute the SVD of a square singular matrix A.

1 Compute $A^{\mathrm{T}}A$ and its eigenvalues $\sigma_1^2, 0$ and unit eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2$:

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}.$$

2 (a) Compute AA^{T} and its eigenvalues $\sigma_1^2, 0$ and unit eigenvectors $\boldsymbol{u}_1, \boldsymbol{u}_2$.

(b) Verify from Problem 1 that $A\boldsymbol{v}_1 = \sigma_1 \boldsymbol{u}_1$. Find all entries in the SVD:

$$\begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}^{\mathrm{T}}.$$

3 Write down orthonormal bases for the four fundamental subspaces of this A.

Problems 4–7 ask for the SVD of matrices of rank 2.

4 (a) Find the eigenvalues and unit eigenvectors of $A^{\mathrm{T}}A$ and AA^{T} for the Fibonacci matrix

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

(b) Construct the singular value decomposition of A.

5 Show that the vectors in Problem 4 satisfy $A\boldsymbol{v}_1 = \sigma_1 \boldsymbol{u}_1$ and $A\boldsymbol{v}_2 = \sigma_2 \boldsymbol{u}_2$.

6 Use the SVD part of the MATLAB demo **eigshow** to find the same vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ graphically.

7 Compute $A^{\mathrm{T}}A$ and AA^{T} and their eigenvalues and unit eigenvectors for

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Multiply the three matrices $U\Sigma V^{\mathrm{T}}$ to recover A.

Problems 8–15 bring out the underlying ideas of the SVD.

8 Suppose $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are orthonormal bases for $\mathbf{R}^n$. Construct the matrix A that transforms each $\boldsymbol{v}_j$ into $\boldsymbol{u}_j$ to give $A\boldsymbol{v}_1 = \boldsymbol{u}_1, \ldots, A\boldsymbol{v}_n = \boldsymbol{u}_n$.

9 Construct the matrix with rank one that has $A\boldsymbol{v} = 12\boldsymbol{u}$ for $\boldsymbol{v} = \frac{1}{2}(1, 1, 1, 1)$ and $\boldsymbol{u} = \frac{1}{3}(2, 2, 1)$.Its only singular value is $\sigma_1 =$ ______.

10 Suppose A has orthogonal columns $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ of lengths $\sigma_1, \sigma_2, \ldots, \sigma_n$. What are U, Σ, and V in the SVD?

11 Explain how the SVD expresses the matrix A as the sum of r rank one matrices:

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}} + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^{\mathrm{T}}.$$

12 Suppose A is a 2 by 2 symmetric matrix with unit eigenvectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. If its eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -2$, what are the matrices U, Σ, V^{T} in its SVD?

13 If $A = QR$ with an orthonormal matrix Q, then the SVD of A is almost the same as the SVD of R. Which of the three matrices in the SVD is changed because of Q?

14 Suppose A is invertible (with $\sigma_1 > \sigma_2 > 0$). Change A by as small a matrix as possible to produce a singular matrix A_0. Hint: U and V do not change:

$$A = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}^{\mathrm{T}}$$

15 (a) If A changes to $4A$, what is the change in the SVD?

(b) What is the SVD for A^{T} and for A^{-1}?

16 Why doesn't the SVD for $A + I$ just use $\Sigma + I$?

17 (MATLAB) Run a random walk starting from web site $x(1) = 1$ and record the visits to each site. From the site $x(k-1) = 1, 2, 3,$ or 4 the code chooses $x(k)$ with probabilities given by column $x(k-1)$ of A. At the end $\boldsymbol{p}$ gives the fraction of time at each site from a histogram (and $A\boldsymbol{p} \approx \boldsymbol{p}$—please check this steady state eigenvector):

$$A = [0\ \ .1\ \ .2\ \ .7;\ \ .05\ \ 0\ \ .15\ \ .8;\ \ .15\ \ .25\ \ 0\ \ .6;\ \ .1\ \ .3\ \ .6\ \ 0]' =$$

Markov matrix

```
n = 1000;  x = zeros(1,n);  x(1) = 1;
for k = 2:n  x(k) = min(find(rand<cumsum(A(:,x(k-1)))));  end
p = hist(x,1:4)/n
```

How are the properties of a matrix reflected in its eigenvalues and eigenvectors? This question is fundamental throughout Chapter 6. A table that organizes the key facts may be helpful. For each class of matrices, here are the special properties of the eigenvalues λ_i and eigenvectors $\boldsymbol{x}_i$.

Symmetric:		
$A^{\mathrm{T}} = A$	real λ's	orthogonal $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j = 0$
Orthogonal:		
$Q^{\mathrm{T}} = Q^{-1}$	all $\lvert\lambda\rvert = 1$	orthogonal $\overline{\boldsymbol{x}}_i^{\mathrm{T}}\boldsymbol{x}_j = 0$
Skew-symmetric:		
$A^{\mathrm{T}} = -A$	imaginary λ's	orthogonal $\overline{\boldsymbol{x}}_i^{\mathrm{T}}\boldsymbol{x}_j = 0$
Complex Hermitian:		
$\overline{A}^{\mathrm{T}} = A$	real λ's	orthogonal $\overline{\boldsymbol{x}}_i^{\mathrm{T}}\boldsymbol{x}_j = 0$
Positive Definite:		
$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} > 0$	all $\lambda > 0$	orthogonal
Markov:		
$m_{ij} > 0, \sum_{i=1}^{n} m_{ij} = 1$	$\lambda_{\max} = 1$	steady state $\boldsymbol{x} > 0$
Similar:		
$B = M^{-1}AM$	$\lambda(B) = \lambda(A)$	$\boldsymbol{x}(B) = M^{-1}\boldsymbol{x}(A)$
Projection:		
$P = P^2 = P^{\mathrm{T}}$	$\lambda = 1;\ 0$	column space; nullspace
Reflection:		
$I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$	$\lambda = -1;\ 1, .., 1$	$\boldsymbol{u}; \boldsymbol{u}^{\perp}$
Rank One:		
$\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$	$\lambda = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u};\ 0, .., 0$	$\boldsymbol{u};\ \boldsymbol{v}^{\perp}$
Inverse:		
A^{-1}	$1/\lambda(A)$	eigenvectors of A
Shift:		
$A + cI$	$\lambda(A) + c$	eigenvectors of A
Stable Powers:		
$A^n \to 0$	all $\lvert\lambda\rvert < 1$	
Stable Exponential:		
$e^{At} \to 0$	all $Re\,\lambda < 0$	
Cyclic Permutation:		
$P(1, .., n) = (2, .., n, 1)$	$\lambda_k = e^{2\pi ik/n}$	$\boldsymbol{x}_k = (1, \lambda_k, \ldots, \lambda_k^{n-1})$
Tridiagonal:		
$-1, 2, -1$ on diagonals	$\lambda_k = 2 - 2\cos\dfrac{k\pi}{n+1}$	$\boldsymbol{x}_k = \left(\sin\dfrac{k\pi}{n+1}, \sin\dfrac{2k\pi}{n+1}, \ldots\right)$
Diagonalizable:		
$S\Lambda S^{-1}$	diagonal of Λ	columns of S are independent
Symmetric:		
$Q\Lambda Q^{\mathrm{T}}$	diagonal of Λ (real)	columns of Q are orthonormal
Jordan:		
$J = M^{-1}AM$	diagonal of J	each block gives $\boldsymbol{x} = (0, .., 1, .., 0)$
Every Matrix:		
$A = U\Sigma V^{\mathrm{T}}$	$\operatorname{rank}(A) = \operatorname{rank}(\Sigma)$	eigenvectors of $A^{\mathrm{T}}A$, AA^{T} in V, U

CHANGE OF BASIS ■ 7.3

This section returns to one of the fundamental ideas of linear algebra—*a basis for* $\mathbf{R}^n$. We don't intend to change that idea, but we do intend to change the basis. It often happens (and we will give examples) that one basis is especially suitable for a specific problem. By changing to that basis, the vectors and the matrices reveal the information we want. The whole idea of a *transform* (this book explains the Fourier transform and wavelet transform) is exactly a change of basis.

Remember what it means for the vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ to be a basis for $\mathbf{R}^n$:

1. The $\boldsymbol{w}$'s are linearly independent.
2. The $n \times n$ matrix W with these columns is invertible.
3. Every vector $\boldsymbol{v}$ in $\mathbf{R}^n$ can be written in exactly one way as a combination of the $\boldsymbol{w}$'s:

$$\boldsymbol{v} = c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 + \cdots + c_n\boldsymbol{w}_n. \tag{1}$$

Here is the key point: Those coefficients $c_1, \ldots, c_n$ completely describe the vector $\boldsymbol{v}$, ***after we have decided on the basis.*** Originally, a column vector $\boldsymbol{v}$ just has the components $v_1, \ldots, v_n$. In the new basis of $\boldsymbol{w}$'s, the same vector is described by the different set of numbers $c_1, \ldots c_n$. It takes n numbers to describe each vector and it also requires a choice of basis. The n numbers are the coordinates of $\boldsymbol{v}$ in that basis:

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}_{\text{standard basis}} \quad \text{and also} \quad \boldsymbol{v} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}_{\text{basis of } \boldsymbol{w}\text{'s}} \tag{2}$$

A basis is a set of axes for $\mathbf{R}^n$. The coordinates $c_1, \ldots, c_n$ tell how far to go along each axis. The axes are at right angles when the $\boldsymbol{w}$'s are orthogonal.

Small point: What is the "*standard basis*"? Those basis vectors are simply the columns of the n by n identity matrix I. These columns $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ are the "default basis." When I write down the vector $\boldsymbol{v} = (2, 4, 5)$ in $\mathbf{R}^3$, I am intending and you are expecting the standard basis (the usual xyz axes, where the coordinates are $2, 4, 5$):

$$\boldsymbol{v} = 2\boldsymbol{e}_1 + 4\boldsymbol{e}_2 + 5\boldsymbol{e}_3 = 2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 5\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}.$$

The new question is: *What are the coordinates* c_1, c_2, c_3 *in the new basis* $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$? As usual we put the basis vectors into the columns of a matrix. This is the *basis matrix* W. Then the fundamental equation $\boldsymbol{v} = c_1\boldsymbol{w}_1 + \ldots + c_n\boldsymbol{w}_n$ has the matrix form $\boldsymbol{v} = W\boldsymbol{c}$. From this we immediately know $\boldsymbol{c} = W^{-1}\boldsymbol{v}$.

7C The coordinates $\boldsymbol{c} = (c_1, \ldots, c_n)$ of $\boldsymbol{v}$ in the basis $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ are given by $\boldsymbol{c} = W^{-1}\boldsymbol{v}$. ***The change of basis matrix*** W^{-1} ***is the inverse of the basis matrix*** W.

The standard basis has $W = I$. The coordinates in that default basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ are the usual components $v_1, \ldots, v_n$. Our first new example is the wavelet basis for $\mathbf{R}^4$.

Example 1 **(Wavelet basis)** Wavelets are little waves. They have different lengths and they are localized at different places. The first basis vector is not actually a wavelet, it is the very useful flat vector of all ones. The others are "Haar wavelets":

$$\boldsymbol{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \boldsymbol{w}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \quad \boldsymbol{w}_3 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \quad \boldsymbol{w}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}. \tag{3}$$

Those vectors are *orthogonal*, which is good. You see how $\boldsymbol{w}_3$ is localized in the first half and $\boldsymbol{w}_4$ is localized in the second half. Their coefficients c_3 and c_4 tell us about details in the first half and last half of $\boldsymbol{v}$. The ultimate in localization is the standard basis.

Why do want to change the basis? I think of v_1, v_2, v_3, v_4 as the intensities of a signal. It could be an audio signal, like music on a CD. It could be a medical signal, like an electrocardiogram. Of course $n = 4$ is very short, and $n = 10,000$ is more realistic. We may need to *compress* that long signal, by keeping only the largest 5% of the coefficients. This is 20 : 1 compression and (to give only one of its applications) it makes modern video conferencing possible.

If we keep only 5% of the *standard* basis coefficients, we lose 95% of the signal. In image processing, most of the image disappears. In audio, 95% of the tape goes blank. But if we choose a better basis of $\boldsymbol{w}$'s, 5% of the basis vectors can come very close to the original signal. In image processing and audio coding, you can't see or hear the difference. We don't need the other 95%!

One good basis vector is a flat $(1, 1, 1, 1)$. That part alone can represent the constant background of our image. A short wave like $(0, 0, 1, -1)$ or in higher dimensions $(0, 0, 0, 0, 0, 0, 1, -1)$ represents a detail at the end of the signal.

The three steps of ***transform*** and ***compression*** and ***inverse transform*** are

$$\begin{array}{ccccccc} \text{input } \boldsymbol{v} & \rightarrow & \text{coefficients } \boldsymbol{c} & \rightarrow & \text{compressed } \widehat{\boldsymbol{c}} & \rightarrow & \text{compressed } \widehat{\boldsymbol{v}} \\ & [\textit{\textbf{lossless}}] & & [\textit{\textbf{lossy}}] & & [\textit{\textbf{reconstruct}}] & \end{array}$$

In linear algebra, where everything is perfect, we omit the compression step. The output $\widehat{\boldsymbol{v}}$ is exactly the same as the input $\boldsymbol{v}$. The transform gives $\boldsymbol{c} = W^{-1}\boldsymbol{v}$ and the reconstruction brings back $\boldsymbol{v} = W\boldsymbol{c}$. In true signal processing, where nothing is perfect but everything is fast, the transform (lossless) and the compression (which only loses unnecessary information) are absolutely the keys to success. Then $\widehat{\boldsymbol{v}} = W\widehat{\boldsymbol{c}}$.

I will show those steps for a typical vector like $\boldsymbol{v} = (6, 4, 5, 1)$. Its wavelet coefficients are $4, 1, 1, 2$. This means that $\boldsymbol{v}$ can be reconstructed from $\boldsymbol{c} = (4, 1, 1, 2)$ using $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \boldsymbol{w}_4$. In matrix form the reconstruction is $\boldsymbol{v} = W\boldsymbol{c}$:

$$\begin{bmatrix} 6 \\ 4 \\ 5 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & -1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 1 \\ 2 \end{bmatrix}. \tag{4}$$

Those coefficients $\boldsymbol{c} = (4, 1, 1, 2)$ are $W^{-1}\boldsymbol{v}$. Inverting this basis matrix W is easy because the $\boldsymbol{w}$'s in its columns are orthogonal. But they are not unit vectors. So the inverse is the transpose divided by the lengths squared, $W^{-1} = (W^{\mathrm{T}}W)^{-1}W^{\mathrm{T}}$:

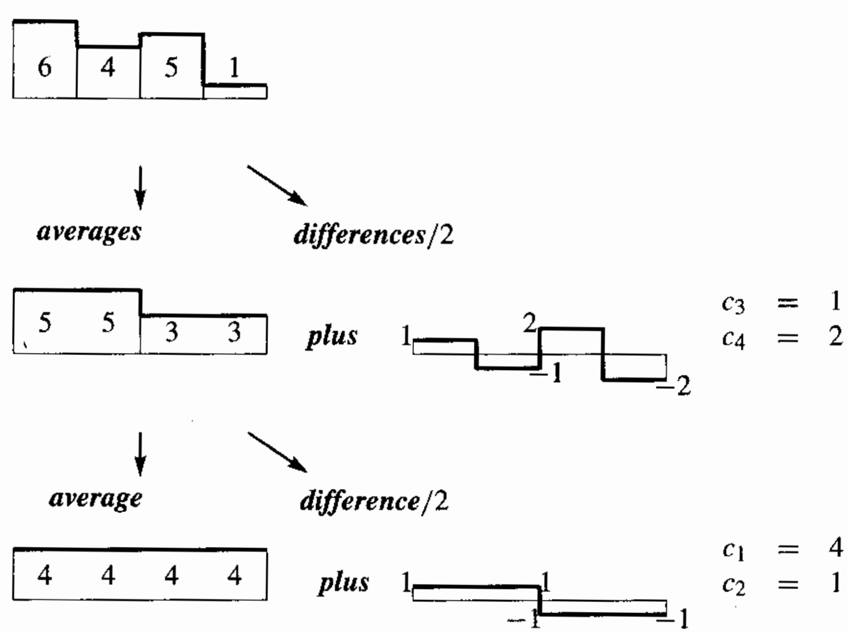
$$W^{-1} = \begin{bmatrix} \frac{1}{4} & & & \\ & \frac{1}{4} & & \\ & & \frac{1}{2} & \\ & & & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

From the 1's in the first row of $\boldsymbol{c} = W^{-1}\boldsymbol{v}$, notice that c_1 is the average of v_1, v_2, v_3, v_4:

$$c_1 = \frac{6+4+5+1}{4} = 4.$$

Example 2 **(Same wavelet basis by recursion)** I can't resist showing you a faster way to find the c's. The special point of the wavelet basis is that you can pick off the details in c_3 and c_4, before the coarse details in c_2 and the overall average in c_1. A picture will explain this "multiscale" method, which is in Chapter 1 of my textbook with Nguyen on *Wavelets and Filter Banks*:

Split $\boldsymbol{v} = (6, 4, 5, 1)$ ***into averages and waves at small scale and then large scale:***



Example 3 **(Fourier basis)** The first thing an electrical engineer does with a signal is to take its Fourier transform. This is a discrete signal (a vector v) and we are speaking about its ***Discrete Fourier Transform***. The DFT involves complex numbers. But if we choose $n = 4$, the matrices are small and the only complex numbers are i and i^3.

Notice that $i^3 = -i$ because $i^2 = -1$. A true electrical engineer would write j instead of i. The four basis vectors are in the columns of the Fourier matrix F:

$$F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix}.$$

The first column is the useful flat basis vector $(1, 1, 1, 1)$. It represents the average signal or the direct current (the DC term). It is a wave at zero frequency. The third column is $(1, -1, 1, -1)$, which alternates at the highest frequency. *The Fourier transform decomposes the signal into waves at equally spaced frequencies.*

The Fourier matrix F is absolutely the most important complex matrix in mathematics and science and engineering. The last section of this book explains the ***Fast Fourier Transform***: it is a factorization of F into matrices with many zeros. The FFT has revolutionized entire industries, by speeding up the Fourier transform. The beautiful thing is that F^{-1} looks like F, with i changed to $-i$:

$$F^{-1} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & (-i) & (-i)^2 & (-i)^3 \\ 1 & (-i)^2 & (-i)^4 & (-i)^6 \\ 1 & (-i)^3 & (-i)^6 & (-i)^9 \end{bmatrix} = \frac{1}{4}\overline{F}.$$

The MATLAB command $c = \text{fft}(v)$ produces the Fourier coefficients $c_1, \ldots, c_n$ of the vector v. It multiplies v by F^{-1} (fast).

The Dual Basis

The columns of W contain the basis vectors $w_1, \ldots, w_n$. To find the coefficients $c_1, \ldots, c_n$ of a vector in this basis, we use the matrix W^{-1}. This subsection just introduces a notation and a new word for the rows of W^{-1}. The vectors in those rows (call them $u_1^{\mathrm{T}}, \ldots, u_n^{\mathrm{T}}$) are the *dual basis.*

The properties of the dual basis reflect $W^{-1}W = I$ and also $WW^{-1} = I$. The product $W^{-1}W$ takes rows of W^{-1} times columns of W, in other words dot products of the u's with the w's. The two bases are "biorthogonal" because we get 1's and 0's:

$$W^{-1}W = \begin{bmatrix} u_1^{\mathrm{T}} \\ \vdots \\ u_n^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} = I \quad \text{so} \quad u_i^{\mathrm{T}} w_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

For an orthonormal basis, the u's are the same as the w's. We have been calling them q's. The basis of q's is biorthogonal to itself! The rows in W^{-1} are the same as the columns in W. In other words $W^{-1} = W^{\mathrm{T}}$. That is the specially important case of an *orthogonal matrix.*

Other bases are not orthonormal. The axes don't have to be perpendicular. The basis matrix W can be invertible without having orthogonal columns.

When the inverse matrices are in the opposite order $WW^{-1} = I$, we learn something new. The columns are $\boldsymbol{w}_j$, the rows are $\boldsymbol{u}_i^{\mathrm{T}}$, and each product is a rank one matrix. *Multiply columns times rows*:

$$WW^{-1} = \begin{bmatrix} \boldsymbol{w}_1 \cdots \boldsymbol{w}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{u}_n^{\mathrm{T}} \end{bmatrix} = \boldsymbol{w}_1\boldsymbol{u}_1^{\mathrm{T}} + \cdots + \boldsymbol{w}_n\boldsymbol{u}_n^{\mathrm{T}} = I.$$

WW^{-1} is the order that we constantly use to change the basis. The coefficients are in $\boldsymbol{c} = W^{-1}\boldsymbol{v}$. So W^{-1} is the first (with the $\boldsymbol{u}_i^{\mathrm{T}}$ in its rows). Then we reconstruct $\boldsymbol{v}$ from $W\boldsymbol{c}$. Use the $\boldsymbol{u}$'s and $\boldsymbol{w}$'s to state the basic facts that $\boldsymbol{c} = W^{-1}\boldsymbol{v}$ and $\boldsymbol{v} = W\boldsymbol{c} = WW^{-1}\boldsymbol{v}$:

$$\text{The coefficients are} \quad c_i = \boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{v} \quad \text{and the vector is} \quad \boldsymbol{v} = \sum_1^n \boldsymbol{w}_i(\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{v}). \tag{5}$$

The analysis step takes dot products of $\boldsymbol{v}$ with the dual basis to find the c's. The synthesis step adds up the pieces $c_i\boldsymbol{w}_i$ to reconstruct the vector $\boldsymbol{v}$.

■ REVIEW OF THE KEY IDEAS ■

1. The new basis vectors $\boldsymbol{w}_j$ are the columns of an invertible matrix W.
2. The coefficients of $\boldsymbol{v}$ in this new basis are $\boldsymbol{c} = W^{-1}\boldsymbol{v}$ (the analysis step).
3. The vector $\boldsymbol{v}$ is reconstructed as $W\boldsymbol{c} = c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n$ (the synthesis step).
4. Compression would simplify $\boldsymbol{c}$ to $\widehat{\boldsymbol{c}}$ and we reconstruct $\widehat{\boldsymbol{v}} = \widehat{c}_1\boldsymbol{w}_1 + \cdots + \widehat{c}_n\boldsymbol{w}_n$.
5. The rows of W^{-1} are the dual basis vectors $\boldsymbol{u}_i$ and $c_i = \boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{v}$. Then $\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{w}_j = \delta_{ij}$.

■ WORKED EXAMPLES ■

7.3 A Match $a_0 + a_1x + a_2x^2$ with $b_0 + b_1(x+1) + b_2(x+1)^2$, to find the 3 by 3 matrix M_1 that connects these coefficients by $\boldsymbol{a} = M_1\boldsymbol{b}$. M_1 will be familiar to Pascal!

The matrix to reverse that change must be M_1^{-1}, and $\boldsymbol{b} = M_1^{-1}\boldsymbol{a}$. This shifts the center of the series back, so $a_0 + a_1(x-1) + a_2(x-1)^2$ equals $b_0 + b_1x + b_2x^2$. Match those quadratics to find M_{-1}, the inverse of Pascal. Also find M_t from $a_0 + a_1x + a_2x^2 = b_0 + b_1(x+t) + b_2(x+t)^2$. Verify that $M_sM_t = M_{s+t}$.

Solution Match $a_0 + a_1x + a_2x$ with $b_0 + b_1(x+1) + b_2(x+1)^2$ to find M_1:

$$\begin{array}{lll} \text{Constant term2} & a_0 = & b_0 + b_1 + \ b_2 \\ \text{Coefficient of } x2 & a_1 = & \qquad b_1 + 2b_2 \\ \text{Coefficient of } x^22 & a_2 = & \qquad\qquad b_2 \end{array} \qquad \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 2 \\ & & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$$

By writing $(x+1)^2 = 1 + 2x + x^2$ we see $1, 2, 1$ in this change of basis matrix.

The matrix M_1 is Pascal's upper triangular P_U. Its inverse M_1^{-1} comes by matching $a_0 + a_1(x-1) + a_2(x-1)^2$ with $b_0 + b_1x + b_2x^2$. The constant terms are the same if $a_0 - a_1 + a_2 = b_0$. This gives alternating signs in $M_1^{-1} = M_{-1}$.

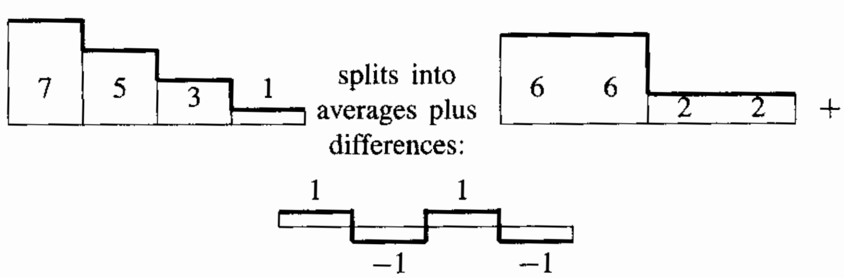
$$\textbf{Inverse of } \mathbf{M_1} = M_{-1} = \begin{bmatrix} 1 & -1 & 1 \\ & 1 & -2 \\ & & 1 \end{bmatrix} \qquad \textbf{Shift by t} \quad M_t = \begin{bmatrix} 1 & t & t^2 \\ & 1 & 2t \\ & & 1 \end{bmatrix}.$$

$M_sM_t = M_{s+t}$ and $M_1M_{-1} = M_0 = I$. Pascal fans might wonder if his symmetric matrix P_S also appears in a change of basis. It does, when the new basis has negative powers $(x+1)^{-k}$ (more about this on the course website **web.mit.edu/18.06/www**).

Problem Set 7.3

1 Express the vectors $\boldsymbol{e} = (1,0,0,0)$ and $\boldsymbol{v} = (1,-1,1,-1)$ in the wavelet basis, as in equation (4). The coefficients c_1, c_2, c_3, c_4 solve $W\boldsymbol{c} = \boldsymbol{e}$ and $W\boldsymbol{c} = \boldsymbol{v}$.

2 Follow Example 2 to represent $\boldsymbol{v} = (7,5,3,1)$ in the wavelet basis. Start with



The last step writes 6, 6, 2, 2 as an overall average plus a difference, using 1, 1, 1, 1 and 1, 1, −1, −1.

3 What are the eight vectors in the wavelet basis for $\mathbf{R}^8$? They include the long wavelet $(1,1,1,1,-1,-1,-1,-1)$ and the short wavelet $(1,-1,0,0,0,0,0,0)$.

4 The wavelet basis matrix W factors into simpler matrices W_1 and W_2:

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

Then $W^{-1} = W_2^{-1}W_1^{-1}$ allows $\boldsymbol{c}$ to be computed in two steps. The first splitting in Example 2 shows $W_1^{-1}\boldsymbol{v}$. Then the second splitting applies W_2^{-1}. Find those inverse matrices W_1^{-1} and W_2^{-1} directly from W_1 and W_2. Apply them to $\boldsymbol{v} = (6, 4, 5, 1)$.

5 The 4 by 4 *Hadamard matrix* is like the wavelet matrix but entirely $+1$ and -1:

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Find H^{-1} and write $\boldsymbol{v} = (7, 5, 3, 1)$ as a combination of the columns of H.

6 Suppose we have two bases $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ for $\mathbf{R}^n$. If a vector has coefficients b_i in one basis and c_i in the other basis, what is the change of basis matrix in $\boldsymbol{b} = M\boldsymbol{c}$? Start from

$$b_1\boldsymbol{v}_1 + \cdots + b_n\boldsymbol{v}_n = V\boldsymbol{b} = c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n = W\boldsymbol{c}.$$

Your answer represents $T(\boldsymbol{v}) = \boldsymbol{v}$ with input basis of $\boldsymbol{v}$'s and output basis of $\boldsymbol{w}$'s. Because of different bases, the matrix is not I.

7 The dual basis vectors $\boldsymbol{w}_1^*, \ldots, \boldsymbol{w}_n^*$ are the columns of $W^* = (W^{-1})^{\mathrm{T}}$. Show that the original basis $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ is "the dual of the dual." In other words, show that the $\boldsymbol{w}$'s are the rows of $(W^*)^{-1}$. **Hint:** Transpose the equation $WW^{-1} = I$.

DIAGONALIZATION AND THE PSEUDOINVERSE ■ 7.4

This short section combines the ideas from Section 7.2 (matrix of a linear transformation) and Section 7.3 (change of basis). The combination produces a needed result: *the change of matrix due to change of basis.* The matrix depends on the input basis and output basis. We want to produce a better matrix than A, by choosing a better basis than the standard basis.

By reversing the input and output bases, we will find the pseudoinverse A^+. It sends $\mathbf{R}^m$ back to $\mathbf{R}^n$, column space back to row space.

The truth is that all our great factorizations of A can be regarded as a change of basis. But this is a short section, so we concentrate on the two outstanding examples. In both cases the good matrix is *diagonal.* It is either Λ or Σ:

1. $S^{-1}AS = \Lambda$ *when the input and output bases are eigenvectors of* A.

2. $U^{-1}AV = \Sigma$ *when the input and output bases are eigenvectors of* $A^{\mathrm{T}}A$ *and* AA^{T}.

You see immediately the difference between Λ and Σ. In Λ the bases are the same. The matrix A must be square. And some square matrices cannot be diagonalized by any S, because they don't have n independent eigenvectors.

In Σ the input and output bases are different. The matrix A can be rectangular. The bases are *orthonormal* because $A^{\mathrm{T}}A$ and AA^{T} are symmetric. Then $U^{-1} = U^{\mathrm{T}}$ and $V^{-1} = V^{\mathrm{T}}$. Every matrix A is allowed, and can be diagonalized. This is the Singular Value Decomposition (**SVD**) of Section 6.7.

I will just note that the Gram-Schmidt factorization $A = QR$ chooses only *one* new basis. That is the orthogonal output basis given by Q. The input uses the standard basis given by I. We don't reach a diagonal Σ, but we do reach a triangular R. The output basis matrix appears on the left and the input basis appears on the right, in $A = QRI$.

We start with input basis equal to output basis. That will produce S and S^{-1}.

Similar Matrices: A and $S^{-1}AS$ and $W^{-1}AW$

We begin with a square matrix and one basis. The input space $\mathbf{V}$ is $\mathbf{R}^n$ and the output space $\mathbf{W}$ is also $\mathbf{R}^n$. The standard basis vectors are the columns of I. The matrix is n by n, and we call it A. The linear transformation T is "multiplication by A".

Most of this book has been about one fundamental problem—*to make the matrix simple.* We made it triangular in Chapter 2 (by elimination) and Chapter 4 (by Gram-Schmidt). We made it diagonal in Chapter 6 (*by eigenvectors*). Now that change from A to Λ comes from a ***change of basis.***

Here are the main facts in advance. When you change the basis for $\mathbf{V}$, the matrix changes from A to AM. Because $\mathbf{V}$ is the input space, the matrix M goes on the right (to come first). When you change the basis for $\mathbf{W}$, the new matrix is $M^{-1}A$. We are working with the output space so M^{-1} is on the left (to come last). ***If you change both bases in the same way, the new matrix is*** $M^{-1}AM$. The good basis vectors are the eigenvectors of A, in the columns of $M = S$. The matrix becomes $S^{-1}AS = \Lambda$.

7D When the basis contains the eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, the matrix for T becomes Λ.

Reason To find column 1 of the matrix, input the first basis vector $\boldsymbol{x}_1$. The transformation multiplies by A. The output is $A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$. This is λ_1 times the first basis vector plus zero times the other basis vectors. Therefore the first column of the matrix is $(\lambda_1, 0, \ldots, 0)$. ***In the eigenvector basis, the matrix is diagonal.***

Example 1 Find the diagonal matrix that projects onto the 135° line $y = -x$. The standard basis $(1, 0)$ and $(0, 1)$ is projected to $(.5, -.5)$ and $(-.5, .5)$

$$\textbf{Standard matrix} \qquad A = \begin{bmatrix} .5 & -.5 \\ -.5 & .5 \end{bmatrix}.$$

Solution The eigenvectors for this projection are $\boldsymbol{x}_1 = (1, -1)$ and $\boldsymbol{x}_2 = (1, 1)$. The first eigenvector lies on the 135° line and the second is perpendicular.
Their projections are $\boldsymbol{x}_1$ and $\boldsymbol{0}$. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 0$. In the eigenvector basis, $P\boldsymbol{x}_1 = \boldsymbol{x}_1$ and $P\boldsymbol{x}_2 = \boldsymbol{0}$ go into the columns of Λ:

$$\textbf{Diagonalized matrix} \qquad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

What if you choose another basis like $\boldsymbol{v}_1 = \boldsymbol{w}_1 = (2, 0)$ and $\boldsymbol{v}_2 = \boldsymbol{w}_2 = (1, 1)$? Since $\boldsymbol{w}_1$ is not an eigenvector, the matrix B in this basis will not be diagonal. The first way to compute B follows the rule of Section 7.2: Find column j of the matrix by writing the output $A\boldsymbol{v}_j$ as a combination of $\boldsymbol{w}$'s.

Apply the projection T to $(2, 0)$. The result is $(1, -1)$ which is $\boldsymbol{w}_1 - \boldsymbol{w}_2$. So the first column of B contains 1 and -1. The second vector $\boldsymbol{w}_2 = (1, 1)$ projects to zero, so the second column of B contains 0 and 0:

$$\text{The matrix is } B = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \text{ in the basis } \boldsymbol{w}_1, \boldsymbol{w}_2. \tag{1}$$

The second way to find the same B is more insightful. Use W^{-1} and W to change between the standard basis and the basis of $\boldsymbol{w}$'s. Those change of basis matrices from Section 7.3 are representing the identity transformation. The product of transformations is just ITI, and the product of matrices is $B = W^{-1}AW$. *B is **similar** to A.*

7E For any basis $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ find the matrix B in three steps. Change the input basis to the standard basis with W. The matrix in the standard basis is A. Then change the output basis back to the $\boldsymbol{w}$'s with W^{-1}. The product $B = W^{-1}AW$ represents ITI:

$$B_{\boldsymbol{w}\text{'s to }\boldsymbol{w}\text{'s}} = W^{-1}_{\text{standard to }\boldsymbol{w}\text{'s}} \quad A_{\text{standard}} \quad W_{\boldsymbol{w}\text{'s to standard}} \tag{2}$$

Example 2 (continuing with the projection) Apply this $W^{-1}AW$ rule to find B, when the basis $(2,0)$ and $(1,1)$ is in the columns of W:

$$W^{-1}AW = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}.$$

The $W^{-1}AW$ rule has produced the same B as in equation (1). ***A change of basis produces a similarity transformation in the matrix.*** The matrices A and B are similar. They have the same eigenvalues (1 and 0). And Λ is similar too.

The Singular Value Decomposition (SVD)

Now the input basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ can be different from the output basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$. In fact the input space $\mathbf{R}^n$ can be different from the output space $\mathbf{R}^m$. Again the best matrix is diagonal (now m by n). To achieve this diagonal matrix Σ, each input vector $\boldsymbol{v}_j$ must transform into a multiple of the output vector $\boldsymbol{u}_j$. That multiple is the *singular value* σ_j on the main diagonal of Σ:

$$\textbf{SVD} \qquad A\boldsymbol{v}_j = \begin{cases} \sigma_j \boldsymbol{u}_j & \text{for } j \le r \\ \mathbf{0} & \text{for } j > r \end{cases} \quad \text{with orthonormal bases.} \tag{3}$$

The singular values are in the order $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r$. The rank r enters because (by definition) singular values are not zero. The second part of the equation says that $\boldsymbol{v}_j$ is in the nullspace for $j = r+1, \ldots, n$. This gives the correct number $n - r$ of basis vectors for the nullspace.

Let me connect the matrices A and Σ and V and U with the linear transformations they represent. The matrices A and Σ represent the *same transformation*. $A = U\Sigma V^{\mathrm{T}}$ uses the standard bases for $\mathbf{R}^n$ and $\mathbf{R}^m$. The diagonal Σ uses the input basis of $\boldsymbol{v}$'s and the output basis of $\boldsymbol{u}$'s. The orthogonal matrices V and U give the basis changes; they represent the identity transformations (in $\mathbf{R}^n$ and $\mathbf{R}^m$). The product of transformations is ITI, and it is represented in the $\boldsymbol{v}$ and $\boldsymbol{u}$ bases by $U^{-1}AV$ which is Σ:

7F The matrix Σ in the new bases comes from A in the standard bases by $U^{-1}AV$:

$$\Sigma_{\boldsymbol{v}\text{'s to }\boldsymbol{u}\text{'s}} = U^{-1}_{\text{standard to }\boldsymbol{u}\text{'s}} \; A_{\text{standard}} \; V_{\boldsymbol{v}\text{'s to standard}}. \tag{4}$$

The SVD chooses orthonormal bases ($U^{-1} = U^{\mathrm{T}}$ and $V^{-1} = V^{\mathrm{T}}$) that diagonalize A.

The two orthonormal bases in the SVD are the eigenvector bases for $A^{\mathrm{T}}A$ (the $\boldsymbol{v}$'s) and AA^{T} (the $\boldsymbol{u}$'s). Since those are symmetric matrices, their unit eigenvectors are orthonormal. Their eigenvalues are the numbers σ_j^2. Equations (10) and (11) in Section 6.7 proved that those bases diagonalize the standard matrix A to produce Σ.

Polar Decomposition

Every complex number has the polar form $re^{i\theta}$. A nonnegative number r multiplies a number on the unit circle. (Remember that $|e^{i\theta}| = |\cos\theta + i\sin\theta| = 1$.) Thinking of these numbers as 1 by 1 matrices, $r \geq 0$ corresponds to a *positive semidefinite matrix* (call it H) and $e^{i\theta}$ corresponds to an *orthogonal matrix* Q. The ***polar decomposition*** extends this $re^{i\theta}$ factorization to matrices.

7G Every real square matrix can be factored into $A = QH$, where Q is ***orthogonal*** and H is ***symmetric positive semidefinite***. If A is invertible then H is positive definite.

For the proof we just insert $V^{\mathrm{T}}V = I$ into the middle of the SVD:

$$A = U\Sigma V^{\mathrm{T}} = (UV^{\mathrm{T}})(V\Sigma V^{\mathrm{T}}) = (Q)(H). \tag{5}$$

The first factor UV^{T} is Q. The product of orthogonal matrices is orthogonal. The second factor $V\Sigma V^{\mathrm{T}}$ is H. It is positive semidefinite because its eigenvalues are in Σ. If A is invertible then Σ and H are also invertible. ***H is the symmetric positive definite square root of $A^{\mathrm{T}}A$.*** Equation (5) says that $H^2 = V\Sigma^2V^{\mathrm{T}} = A^{\mathrm{T}}A$.

There is also a polar decomposition $A = KQ$ in the reverse order. Q is the same but now $K = U\Sigma U^{\mathrm{T}}$. This is the symmetric positive definite square root of AA^{T}.

Example 3 Find the polar decomposition $A = QH$ from its SVD in Section 6.7:

$$A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2} & \\ & 2\sqrt{2} \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = U\Sigma V^{\mathrm{T}}.$$

Solution The orthogonal part is $Q = UV^{\mathrm{T}}$. The positive definite part is $H = V\Sigma V^{\mathrm{T}}$. This is also $H = Q^{-1}A$ which is $Q^{\mathrm{T}}A$:

$$Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$$H = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 3/\sqrt{2} \end{bmatrix}.$$

In mechanics, the polar decomposition separates the ***rotation*** (in Q) from the ***stretching*** (in H). The eigenvalues of H are the singular values of A. They give the stretching factors. The eigenvectors of H are the eigenvectors of $A^{\mathrm{T}}A$. They give the stretching directions (the principal axes). Then Q rotates the axes.

The polar decomposition just splits the key equation $A\boldsymbol{v}_i = \sigma_i\boldsymbol{u}_i$ into two steps. The "H" part multiplies $\boldsymbol{v}_i$ by σ_i. The "Q" part swings $\boldsymbol{v}_i$ around into $\boldsymbol{u}_i$.

The Pseudoinverse

By choosing good bases, A multiplies $\boldsymbol{v}_i$ in the row space to give $\sigma_i \boldsymbol{u}_i$ in the column space. A^{-1} must do the opposite! If $A\boldsymbol{v} = \sigma \boldsymbol{u}$ then $A^{-1}\boldsymbol{u} = \boldsymbol{v}/\sigma$. The singular values of A^{-1} are $1/\sigma$, just as the eigenvalues of A^{-1} are $1/\lambda$. The bases are reversed. The $\boldsymbol{u}$'s are in the row space of A^{-1}, the $\boldsymbol{v}$'s are in the column space.

Until this moment we would have added "*if A^{-1} exists.*" Now we don't. A matrix that multiplies $\boldsymbol{u}_i$ to produce $\boldsymbol{v}_i/\sigma_i$ *does* exist. It is the pseudoinverse A^+:

Pseudoinverse

$$A^+ = V\Sigma^+U^{\mathrm{T}} = \underset{n \text{ by } n}{\begin{bmatrix} \boldsymbol{v}_1 \cdots \boldsymbol{v}_r \cdots \boldsymbol{v}_n \end{bmatrix}} \underset{n \text{ by } m}{\begin{bmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_r^{-1} \\ & & & \end{bmatrix}} \underset{m \text{ by } m}{\begin{bmatrix} \boldsymbol{u}_1 \cdots \boldsymbol{u}_r \cdots \boldsymbol{u}_m \end{bmatrix}^{\mathrm{T}}}$$

The ***pseudoinverse*** A^+ is an n by m matrix. If A^{-1} exists (we said it again), then A^+ is the same as A^{-1}. In that case $m = n = r$ and we are inverting $U\Sigma V^{\mathrm{T}}$ to get $V\Sigma^{-1}U^{\mathrm{T}}$. The new symbol A^+ is needed when $r < m$ or $r < n$. Then A has no two-sided inverse, but it has a *pseudo*inverse A^+ with that same rank r:

$$A^+\boldsymbol{u}_i = \frac{1}{\sigma_i}\boldsymbol{v}_i \quad \text{for } i \le r \qquad \text{and} \qquad A^+\boldsymbol{u}_i = \boldsymbol{0} \quad \text{for } i > r.$$

The vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ in the column space of A go back to the row space. The other vectors $\boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_m$ are in the left nullspace, and A^+ sends them to zero. When we know what happens to each basis vector $\boldsymbol{u}_i$, we know A^+.

Notice the pseudoinverse Σ^+ of the diagonal matrix Σ. Each σ is replaced by σ^{-1}. The product $\Sigma^+\Sigma$ is as near to the identity as we can get. We get r 1's. We can't do anything about the zero rows and columns! This example has $\sigma_1 = 2$ and $\sigma_2 = 3$:

$$\Sigma^+\Sigma = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

7H The pseudoinverse A^+ is the n by m matrix with these two properties:

$$AA^+ = \text{projection matrix onto the column space of } A$$
$$A^+A = \text{projection matrix onto the row space of } A$$

Example 4 Find the pseudoinverse of $A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$. This matrix is not invertible. The rank is 1. The only singular value is $\sqrt{10}$. That is inverted to $1/\sqrt{10}$ in Σ^+:

$$A^+ = V\Sigma^+U^{\mathrm{T}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{10} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}.$$

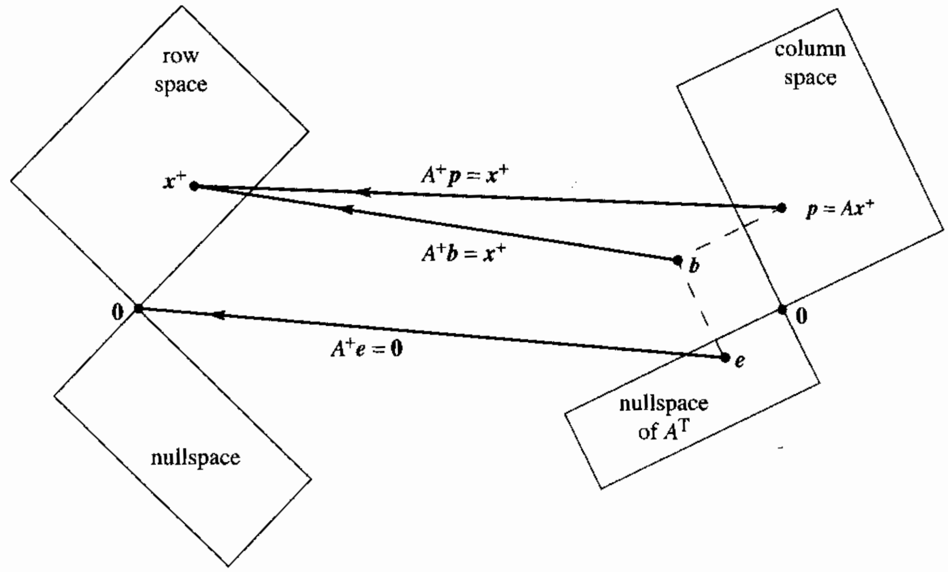


Figure 7.4 A is invertible from row space to column space. A^+ inverts it.

A^+ also has rank 1. Its column space is the row space of A. When A takes $(1, 1)$ in the row space to $(4, 2)$ in the column space, A^+ does the reverse. Every rank one matrix is a column times a row. With unit vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, that is $A = \sigma \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$. Then the best inverse of a rank one matrix is $A^+ = \boldsymbol{v}\boldsymbol{u}^{\mathrm{T}}/\sigma$.

The product AA^+ is $\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$, the projection onto the line through $\boldsymbol{u}$. The product A^+A is $\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$, the projection onto the line through $\boldsymbol{v}$. ***For all matrices,*** AA^+ ***and*** A^+A ***are the projections onto the column space and row space.***

The shortest least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$ is $\boldsymbol{x}^+ = A^+\boldsymbol{b}$. Any other vector that solves $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$ is longer than $\boldsymbol{x}^+$ (Problem 18).

■ REVIEW OF THE KEY IDEAS ■

1. Diagonalization $S^{-1}AS = \Lambda$ is the same as a change to the eigenvector basis.

2. The SVD chooses an input basis of $\boldsymbol{v}$'s and an output basis of $\boldsymbol{u}$'s. Those orthonormal bases diagonalize A. This is $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$, and $A = U\Sigma V^{\mathrm{T}}$.

3. Polar decomposition factors A into QH, rotation times stretching.

4. The pseudoinverse $A^+ = V\Sigma^+U^{\mathrm{T}}$ transforms the column space of A back to its row space. A^+A is the identity on the row space (and zero on the nullspace).

■ WORKED EXAMPLES ■

7.4 A Start with an m by n matrix A. If its rank is n (full column rank) then it has a *left inverse* $C = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$. This matrix C gives $CA = I$. Explain why the pseudoinverse is $A^+ = C$ in this case. If A has rank m (full row rank) then it has a *right inverse* B with $B = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}$ and $AB = I$. Explain why $A^+ = B$ in this case.

Find B and C if possible and find A^+ for all three matrices:

$$A_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad A_2 = \begin{bmatrix} 2 & 2 \end{bmatrix} \quad A_3 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

Solution If A has rank n (independent columns) then $A^{\mathrm{T}}A$ is invertible—this is a key point of Section 4.2. Certainly $C = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ multiplies A to give $CA = I$. In the opposite order, $AC = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ is the projection matrix (Section 4.2 again) onto the column space. So C meets the requirements **7H** to be A^+.

If A has rank m (full row rank) then AA^{T} is invertible. Certainly A multiplies $B = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}$ to give $AB = I$. In the opposite order, $BA = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A$ is the projection matrix onto the row space. So B is the pseudoinverse A^+.

The example A_1 has full column rank (for C) and A_2 has full row rank (for B):

$$A_1^+ = (A_1^{\mathrm{T}}A_1)^{-1}A_1^{\mathrm{T}} = \frac{1}{\sqrt{8}}\begin{bmatrix} 2 & 2 \end{bmatrix} \quad 2A_2^+ = A_2^{\mathrm{T}}(A_2A_2^{\mathrm{T}})^{-1} = \frac{1}{\sqrt{8}}\begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Notice $A_1^+A_1 = [1]$ and $A_2A_2^+ = [1]$. But A_3 has no left or right inverse. Its pseudoinverse is $A_3^+ = \sigma_1^{-1}v_1u_1^{\mathrm{T}} = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]/4$.

Problem Set 7.4

Problems 1–6 compute and use the SVD of a particular matrix (not invertible).

1 Compute $A^{\mathrm{T}}A$ and its eigenvalues and unit eigenvectors v_1 and v_2:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$$

What is the only singular value σ_1? The rank of A is $r = 1$.

2 (a) Compute AA^{T} and its eigenvalues and unit eigenvectors u_1 and u_2.

(b) Verify from Problem 1 that $Av_1 = \sigma_1 u_1$. Put numbers into the SVD:

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix}\begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix}\begin{bmatrix} v_1 & v_2 \end{bmatrix}^{\mathrm{T}}.$$

3 From the $\boldsymbol{u}$'s and $\boldsymbol{v}$'s write down orthonormal bases for the four fundamental subspaces of this matrix A.

4 Describe all matrices that have those same four subspaces.

5 From U, V, and Σ find the orthogonal matrix $Q = UV^{\mathrm{T}}$ and the symmetric matrix $H = V\Sigma V^{\mathrm{T}}$. Verify the polar decomposition $A = QH$. This H is only semidefinite because _____.

6 Compute the pseudoinverse $A^+ = V\Sigma^+ U^{\mathrm{T}}$. The diagonal matrix Σ^+ contains $1/\sigma_1$. Rename the four subspaces (for A) in Figure 7.4 as four subspaces for A^+. Compute A^+A and AA^+.

Problems 7–11 are about the SVD of an invertible matrix.

7 Compute $A^{\mathrm{T}}A$ and its eigenvalues and unit eigenvectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. What are the singular values σ_1 and σ_2 for this matrix A?

$$A = \begin{bmatrix} 3 & 3 \\ -1 & 1 \end{bmatrix}.$$

8 AA^{T} has the same eigenvalues σ_1^2 and σ_2^2 as $A^{\mathrm{T}}A$. Find unit eigenvectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. Put numbers into the SVD:

$$A = \begin{bmatrix} 3 & 3 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}^{\mathrm{T}}.$$

9 In Problem 8, multiply columns times rows to show that $A = \sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}} + \sigma_2\boldsymbol{u}_2\boldsymbol{v}_2^{\mathrm{T}}$. Prove from $A = U\Sigma V^{\mathrm{T}}$ that every matrix of rank r is the sum of r matrices of rank one.

10 From U, V, and Σ find the orthogonal matrix $Q = UV^{\mathrm{T}}$ and the symmetric matrix $K = U\Sigma U^{\mathrm{T}}$. Verify the polar decomposition in the reverse order $A = KQ$.

11 The pseudoinverse of this A is the same as _____ because _____.

Problems 12–13 compute and use the SVD of a 1 by 3 rectangular matrix.

12 Compute $A^{\mathrm{T}}A$ and AA^{T} and their eigenvalues and unit eigenvectors when the matrix is $A = \begin{bmatrix} 3 & 4 & 0 \end{bmatrix}$. What are the singular values of A?

13 Put numbers into the singular value decomposition of A:

$$A = \begin{bmatrix} 3 & 4 & 0 \end{bmatrix} = [\boldsymbol{u}_1]\begin{bmatrix} \sigma_1 & 0 & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{bmatrix}^{\mathrm{T}}.$$

Put numbers into the pseudoinverse of A. *Compute AA^+ and A^+A*:

$$A^+ = \begin{bmatrix} \\ \\ \\ \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{bmatrix} \begin{bmatrix} 1/\sigma_1 \\ 0 \\ 0 \end{bmatrix} [\boldsymbol{u}_1]^{\mathrm{T}}.$$

14 What is the only 2 by 3 matrix that has no pivots and no singular values? What is Σ for that matrix? A^+ is the zero matrix, but what shape?

15 If $\det A = 0$ how do you know that $\det A^+ = 0$?

16 When are the factors in $U\Sigma V^{\mathrm{T}}$ the same as in $Q\Lambda Q^{\mathrm{T}}$? The eigenvalues λ_i must be positive, to equal the σ_i. Then A must be _____ and positive _____.

Problems 17–20 bring out the main properties of A^+ and $x^+ = A^+b$.

17 Suppose all matrices have rank one. The vector $\boldsymbol{b}$ is (b_1, b_2).

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \quad A^{\mathrm{T}} = \begin{bmatrix} .2 & .1 \\ .2 & .1 \end{bmatrix} \quad AA^{\mathrm{T}} = \begin{bmatrix} .8 & .4 \\ .4 & .2 \end{bmatrix} \quad A^{\mathrm{T}}A = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}$$

(a) The equation $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$ has many solutions because $A^{\mathrm{T}}A$ is _____.

(b) Verify that $\boldsymbol{x}^+ = A^+\boldsymbol{b} = (.2b_1 + .1b_2, .2b_1 + .1b_2)$ does solve $A^{\mathrm{T}}A\boldsymbol{x}^+ = A^{\mathrm{T}}\boldsymbol{b}$.

(c) AA^+ projects onto the column space of A. Therefore _____ projects onto the nullspace of A^{T}. Then $A^{\mathrm{T}}(AA^+ - I)\boldsymbol{b} = \boldsymbol{0}$. This gives $A^{\mathrm{T}}A\boldsymbol{x}^+ = A^{\mathrm{T}}\boldsymbol{b}$ and $\widehat{\boldsymbol{x}}$ can be $\boldsymbol{x}^+$.

18 *The vector $\boldsymbol{x}^+$ is the shortest possible solution to $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$.* Reason: The difference $\widehat{\boldsymbol{x}} - \boldsymbol{x}^+$ is in the nullspace of $A^{\mathrm{T}}A$. This is also the nullspace of A. Explain how it follows that

$$\|\widehat{\boldsymbol{x}}\|^2 = \|\boldsymbol{x}^+\|^2 + \|\widehat{\boldsymbol{x}} - \boldsymbol{x}^+\|^2.$$

Any other solution $\widehat{\boldsymbol{x}}$ has greater length than $\boldsymbol{x}^+$.

19 Every $\boldsymbol{b}$ in $\mathbf{R}^m$ is $\boldsymbol{p} + \boldsymbol{e}$. This is the column space part plus the left nullspace part. Every $\boldsymbol{x}$ in $\mathbf{R}^n$ is $\boldsymbol{x}_r + \boldsymbol{x}_n =$ (row space part) + (nullspace part). Then

$$AA^+\boldsymbol{p} = \_\_\_\_ \qquad AA^+\boldsymbol{e} = \_\_\_\_ \qquad A^+A\boldsymbol{x}_r = \_\_\_\_ \qquad A^+A\boldsymbol{x}_n = \_\_\_\_$$

20 Find A^+ and A^+A and AA^+ for the 2 by 1 matrix whose SVD is

$$A = \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} .6 & -.8 \\ .8 & .6 \end{bmatrix} \begin{bmatrix} 5 \\ 0 \end{bmatrix} [1].$$

21 A general 2 by 2 matrix A is determined by four numbers. If triangular, it is determined by three. If diagonal, by two. If a rotation, by one. An eigenvector, by one. Check that the total count is four for each factorization of A:

$$LU \quad LDU \quad QR \quad U\Sigma V^{\mathrm{T}} \quad S\Lambda S^{-1}.$$

22 Following Problem 21, check that LDL^{T} and $Q\Lambda Q^{\mathrm{T}}$ are determined by *three* numbers. This is correct because the matrix A is _____.

23 From $A = U\Sigma V^{\mathrm{T}}$ and $A^{+} = V\Sigma^{+}U^{\mathrm{T}}$ explain these splittings into rank 1:

$$A = \sum_{1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}} \qquad A^{+} = \sum_{1}^{r} \frac{\boldsymbol{v}_i \boldsymbol{u}_i^{\mathrm{T}}}{\sigma_i} \qquad A^{+}A = \sum_{1}^{r} \boldsymbol{v}_i \boldsymbol{v}_i^{\mathrm{T}} \qquad AA^{+} = \sum_{1}^{r} \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{T}}$$

24 This problem looks for all matrices A with a given column space in $\mathbf{R}^m$ and a given row space in $\mathbf{R}^n$. Suppose $\widehat{\boldsymbol{u}}_1, \ldots, \widehat{\boldsymbol{u}}_r$ and $\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_r$ are bases for those two spaces. Make them columns of $\widehat{U}$ and $\widehat{V}$. Use $A = U\Sigma V^{\mathrm{T}}$ to show that A has the form $\widehat{U}M\widehat{V}^{\mathrm{T}}$ for an r by r invertible matrix M.

25 A pair of singular vectors $\boldsymbol{v}$ and $\boldsymbol{u}$ will satisfy $A\boldsymbol{v} = \sigma\boldsymbol{u}$ and $A^{\mathrm{T}}\boldsymbol{u} = \sigma\boldsymbol{v}$. This means that the double vector $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}$ is an eigenvector of what symmetric matrix? With what eigenvalue?

GLOSSARY

Adjacency matrix of a graph. Square matrix with $a_{ij} = 1$ when there is an edge from node i to node j; otherwise $a_{ij} = 0$. $A = A^{\mathrm{T}}$ for an undirected graph.

Affine transformation $T(\boldsymbol{v}) = A\boldsymbol{v} + \boldsymbol{v}_0$ = linear transformation plus shift.

Associative Law $(AB)C = A(BC)$. Parentheses can be removed to leave ABC.

Augmented matrix $[\,A \;\; \boldsymbol{b}\,]$. $A\boldsymbol{x} = \boldsymbol{b}$ is solvable when $\boldsymbol{b}$ is in the column space of A; then $[\,A \;\; \boldsymbol{b}\,]$ has the same rank as A. Elimination on $[\,A \;\; \boldsymbol{b}\,]$ keeps equations correct.

Back substitution. Upper triangular systems are solved in reverse order x_n to x_1.

Basis for $\boldsymbol{V}$. Independent vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d$ whose linear combinations give every $\boldsymbol{v}$ in $\boldsymbol{V}$. A vector space has many bases!

Big formula for n by n determinants. Det(A) is a sum of $n!$ terms, one term for each permutation P of the columns. That term is the product $a_{1\alpha} \cdots a_{n\omega}$ down the diagonal of the reordered matrix, times $\det(P) = \pm 1$.

Block matrix. A matrix can be partitioned into matrix blocks, by cuts between rows and/or between columns. **Block multiplication** of AB is allowed if the block shapes permit (the columns of A and rows of B must be in matching blocks).

Cayley-Hamilton Theorem. $p(\lambda) = \det(A - \lambda I)$ has $p(A) =$ *zero matrix.*

Change of basis matrix M. The old basis vectors $\boldsymbol{v}_j$ are combinations $\sum m_{ij}\boldsymbol{w}_i$ of the new basis vectors. The coordinates of $c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n = d_1\boldsymbol{w}_1 + \cdots + d_n\boldsymbol{w}_n$ are related by $\boldsymbol{d} = M\boldsymbol{c}$. (For $n = 2$ set $\boldsymbol{v}_1 = m_{11}\boldsymbol{w}_1 + m_{21}\boldsymbol{w}_2$, $\boldsymbol{v}_2 = m_{12}\boldsymbol{w}_1 + m_{22}\boldsymbol{w}_2$.)

Characteristic equation $\det(A - \lambda I) = 0$. The n roots are the eigenvalues of A.

Cholesky factorization $A = CC^{\mathrm{T}} = (L\sqrt{D})(L\sqrt{D})^{\mathrm{T}}$ for positive definite A.

Circulant matrix C. Constant diagonals wrap around as in cyclic shift S. Every circulant is $c_0 I + c_1 S + \cdots + c_{n-1}S^{n-1}$. $C\boldsymbol{x} =$ **convolution** $\boldsymbol{c} * \boldsymbol{x}$. Eigenvectors in F.

Cofactor C_{ij}. Remove row i and column j; multiply the determinant by $(-1)^{i+j}$.

Column picture of $A\boldsymbol{x} = \boldsymbol{b}$. The vector $\boldsymbol{b}$ becomes a combination of the columns of A. The system is solvable only when $\boldsymbol{b}$ is in the column space $\boldsymbol{C}(A)$.

Column space $\boldsymbol{C}(A)$ = space of all combinations of the columns of A.

Commuting matrices $AB = BA$. If diagonalizable, they share n eigenvectors.

Companion matrix. Put $c_1, \dots, c_n$ in row n and put $n - 1$ 1's along diagonal 1. Then $\det(A - \lambda I) = \pm(c_1 + c_2\lambda + c_3\lambda^2 + \cdots)$.

Complete solution $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n$ to $A\boldsymbol{x} = \boldsymbol{b}$. (Particular $\boldsymbol{x}_p$) + ($\boldsymbol{x}_n$ in nullspace).

Complex conjugate $\overline{z} = a - ib$ for any complex number $z = a + ib$. Then $z\overline{z} = |z|^2$.

Condition number $cond(A) = \kappa(A) = \|A\|\|A^{-1}\| = \sigma_{\max}/\sigma_{\min}$. In $A\boldsymbol{x} = \boldsymbol{b}$, the relative change $\|\delta\boldsymbol{x}\|/\|\boldsymbol{x}\|$ is less than $cond(A)$ times the relative change $\|\delta\boldsymbol{b}\|/\|\boldsymbol{b}\|$. Condition numbers measure the *sensitivity* of the output to change in the input.

Conjugate Gradient Method. A sequence of steps (end of Chapter 9) to solve positive definite $A\boldsymbol{x} = \boldsymbol{b}$ by minimizing $\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{b}$ over growing Krylov subspaces.

Covariance matrix Σ. When random variables x_i have mean = average value = 0, their covariances Σ_{ij} are the averages of $x_i x_j$. With means $\overline{x}_i$, the matrix Σ = mean of $(\boldsymbol{x} - \overline{\boldsymbol{x}})(\boldsymbol{x} - \overline{\boldsymbol{x}})^{\mathrm{T}}$ is positive (semi)definite; it is diagonal if the x_i are independent.

Cramer's Rule for $A\boldsymbol{x} = \boldsymbol{b}$. B_j has $\boldsymbol{b}$ replacing column j of A, and $x_j = |B_j|/|A|$.

Cross product $\boldsymbol{u} \times \boldsymbol{v}$ in $\mathbf{R}^3$. Vector perpendicular to $\boldsymbol{u}$ and $\boldsymbol{v}$, length $\|\boldsymbol{u}\|\|\boldsymbol{v}\||\sin\theta|$ = parallelogram area, computed as the "determinant" of $[\,\boldsymbol{i}\ \ \boldsymbol{j}\ \ \boldsymbol{k}\,;\ u_1\ \ u_2\ \ u_3\,;\ v_1\ \ v_2\ \ v_3\,]$.

Cyclic shift S. Permutation with $s_{21} = 1, s_{32} = 1, \ldots$, finally $s_{1n} = 1$. Its eigenvalues are nth roots $e^{2\pi ik/n}$ of 1; eigenvectors are columns of the Fourier matrix F.

Determinant $|A| = \det(A)$. Defined by $\det I = 1$, sign reversal for row exchange, and linearity in each row. Then $|A| = 0$ when A is singular. Also $|AB| = |A||B|$ and $|A^{-1}| = 1/|A|$ and $|A^{\mathrm{T}}| = |A|$. The big formula for $\det(A)$ has a sum of $n!$ terms, the cofactor formula uses determinants of size $n - 1$, volume of box $= |\det(A)|$.

Diagonal matrix D. $d_{ij} = 0$ if $i \neq j$. **Block-diagonal**: zero outside square blocks D_{ii}.

Diagonalizable matrix A. Must have n independent eigenvectors (in the columns of S; automatic with n different eigenvalues). Then $S^{-1}AS = \Lambda$ = eigenvalue matrix.

Diagonalization $\Lambda = S^{-1}AS$. Λ = eigenvalue matrix and S = eigenvector matrix. A must have n independent eigenvectors to make S invertible. All $A^k = S\Lambda^k S^{-1}$.

Dimension of vector space $\dim(V)$ = number of vectors in any basis for V.

Distributive Law $A(B + C) = AB + AC$. Add then multiply, or multiply then add.

Dot product $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = x_1y_1 + \cdots + x_ny_n$. Complex dot product is $\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{y}$. Perpendicular vectors have zero dot product. $(AB)_{ij}$ = (row i of A)·(column j of B).

Echelon matrix U. The first nonzero entry (the pivot) in each row comes after the pivot in the previous row. All zero rows come last.

Eigenvalue λ **and eigenvector** $\boldsymbol{x}$. $A\boldsymbol{x} = \lambda\boldsymbol{x}$ with $\boldsymbol{x} \neq \boldsymbol{0}$ so $\det(A - \lambda I) = 0$.

Eigshow. Graphical 2 by 2 eigenvalues and singular values (MATLAB or Java).

Elimination. A sequence of row operations that reduces A to an upper triangular U or to the reduced form $R = \mathrm{rref}(A)$. Then $A = LU$ with multipliers ℓ_{ij} in L, or $PA = LU$ with row exchanges in P, or $EA = R$ with an invertible E.

Elimination matrix = Elementary matrix E_{ij}. The identity matrix with an extra $-\ell_{ij}$ in the i, j entry ($i \neq j$). Then $E_{ij}A$ subtracts ℓ_{ij} times row j of A from row i.

Ellipse (or ellipsoid) $x^{\mathrm{T}}Ax = 1$. A must be positive definite; the axes of the ellipse are eigenvectors of A, with lengths $1/\sqrt{\lambda}$. (For $\|x\| = 1$ the vectors $y = Ax$ lie on the ellipse $\|A^{-1}y\|^2 = y^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}y = 1$ displayed by eigshow; axis lengths σ_i.)

Exponential $e^{At} = I + At + (At)^2/2! + \cdots$ has derivative Ae^{At}; $e^{At}u(0)$ solves $u' = Au$.

Factorization $A = LU$. If elimination takes A to U *without row exchanges*, then the lower triangular L with multipliers ℓ_{ij} (and $\ell_{ii} = 1$) brings U back to A.

Fast Fourier Transform (FFT). A factorization of the Fourier matrix F_n into $\ell = \log_2 n$ matrices S_i times a permutation. Each S_i needs only $n/2$ multiplications, so $F_n x$ and $F_n^{-1}c$ can be computed with $n\ell/2$ multiplications. Revolutionary.

Fibonacci numbers $0, 1, 1, 2, 3, 5, \ldots$ satisfy $F_n = F_{n-1} + F_{n-2} = (\lambda_1^n - \lambda_2^n)/(\lambda_1 - \lambda_2)$. Growth rate $\lambda_1 = (1+\sqrt{5})/2$ is the largest eigenvalue of the Fibonacci matrix $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

Four fundamental subspaces of $A = C(A), N(A), C(A^{\mathrm{T}}), N(A^{\mathrm{T}})$.

Fourier matrix F. Entries $F_{jk} = e^{2\pi ijk/n}$ give orthogonal columns $\overline{F}^{\mathrm{T}}F = nI$. Then $y = Fc$ is the (inverse) Discrete Fourier Transform $y_j = \sum c_k e^{2\pi ijk/n}$.

Free columns of A. Columns without pivots; combinations of earlier columns.

Free variable x_i. Column i has no pivot in elimination. We can give the $n-r$ free variables any values, then $Ax = b$ determines the r pivot variables (if solvable!).

Full column rank $r = n$. Independent columns, $N(A) = \{\mathbf{0}\}$, no free variables.

Full row rank $r = m$. Independent rows, at least one solution to $Ax = b$, column space is all of $\mathbf{R}^m$. *Full rank* means full column rank or full row rank.

Fundamental Theorem. The nullspace $N(A)$ and row space $C(A^{\mathrm{T}})$ are orthogonal complements (perpendicular subspaces of $\mathbf{R}^n$ with dimensions r and $n-r$) from $Ax = \mathbf{0}$. Applied to A^{T}, the column space $C(A)$ is the orthogonal complement of $N(A^{\mathrm{T}})$.

Gauss-Jordan method. Invert A by row operations on $[\,A \quad I\,]$ to reach $[\,I \quad A^{-1}\,]$.

Gram-Schmidt orthogonalization $A = QR$. Independent columns in A, orthonormal columns in Q. Each column q_j of Q is a combination of the first j columns of A (and conversely, so R is upper triangular). Convention: $\operatorname{diag}(R) > \mathbf{0}$.

Graph G. Set of n nodes connected pairwise by m edges. A **complete graph** has all $n(n-1)/2$ edges between nodes. A **tree** has only $n-1$ edges and no closed loops. A **directed graph** has a direction arrow specified on each edge.

Hankel matrix H. Constant along each antidiagonal; h_{ij} depends on $i + j$.

Hermitian matrix $A^{\mathrm{H}} = \overline{A}^{\mathrm{T}} = A$. Complex analog of a symmetric matrix: $\overline{a_{ji}} = a_{ij}$.

Hessenberg matrix H. Triangular matrix with one extra nonzero adjacent diagonal.

Hilbert matrix hilb(n). Entries $H_{ij} = 1/(i+j-1) = \int_0^1 x^{i-1}x^{j-1}dx$. Positive definite but extremely small $\lambda_{\min}$ and large condition number.

Hypercube matrix P_L^2. Row $n+1$ counts corners, edges, faces, $\ldots$ of a cube in $\mathbf{R}^n$.

Identity matrix I (or I_n). Diagonal entries $= 1$, off-diagonal entries $= 0$.

Incidence matrix of a directed graph. The m by n edge-node incidence matrix has a row for each edge (node i to node j), with entries -1 and 1 in columns i and j.

Indefinite matrix. A symmetric matrix with eigenvalues of both signs ($+$ and $-$).

Independent vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$. No combination $c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k =$ zero vector unless all $c_i = 0$. If the $\boldsymbol{v}$'s are the columns of A, the only solution to $A\boldsymbol{x} = \boldsymbol{0}$ is $\boldsymbol{x} = \boldsymbol{0}$.

Inverse matrix A^{-1}. Square matrix with $A^{-1}A = I$ and $AA^{-1} = I$. No inverse if $\det A = 0$ and $\operatorname{rank}(A) < n$ and $A\boldsymbol{x} = \boldsymbol{0}$ for a nonzero vector $\boldsymbol{x}$. The inverses of AB and A^{T} are $B^{-1}A^{-1}$ and $(A^{-1})^{\mathrm{T}}$. Cofactor formula $(A^{-1})_{ij} = C_{ji}/\det A$.

Iterative method. A sequence of steps intended to approach the desired solution.

Jordan form $J = M^{-1}AM$. If A has s independent eigenvectors, its "generalized" eigenvector matrix M gives $J = \operatorname{diag}(J_1, \ldots, J_s)$. The block J_k is $\lambda_k I_k + N_k$ where N_k has 1's on diagonal 1. Each block has one eigenvalue λ_k and one eigenvector $(1, 0, \ldots, 0)$.

Kirchhoff's Laws. *Current law*: net current (in minus out) is zero at each node. *Voltage law*: Potential differences (voltage drops) add to zero around any closed loop.

Kronecker product (tensor product) $A \otimes B$. Blocks $a_{ij}B$, eigenvalues $\lambda_p(A)\lambda_q(B)$.

Krylov subspace $K_j(A, \boldsymbol{b})$. The subspace spanned by $\boldsymbol{b}, A\boldsymbol{b}, \ldots, A^{j-1}\boldsymbol{b}$. Numerical methods approximate $A^{-1}\boldsymbol{b}$ by $\boldsymbol{x}_j$ with residual $\boldsymbol{b} - A\boldsymbol{x}_j$ in this subspace. A good basis for K_j requires only multiplication by A at each step.

Least squares solution $\widehat{\boldsymbol{x}}$. The vector $\widehat{\boldsymbol{x}}$ that minimizes the error $\|\boldsymbol{e}\|^2$ solves $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$. Then $\boldsymbol{e} = \boldsymbol{b} - A\widehat{\boldsymbol{x}}$ is orthogonal to all columns of A.

Left inverse A^+. If A has full column rank n, then $A^+ = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ has $A^+A = I_n$.

Left nullspace $N(A^{\mathrm{T}})$. Nullspace of $A^{\mathrm{T}} =$ "left nullspace" of A because $\boldsymbol{y}^{\mathrm{T}}A = \boldsymbol{0}^{\mathrm{T}}$.

Length $\|\boldsymbol{x}\|$. Square root of $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ (Pythagoras in n dimensions).

Linear combination $c\boldsymbol{v} + d\boldsymbol{w}$ or $\sum c_j\boldsymbol{v}_j$. Vector addition and scalar multiplication.

Linear transformation T. Each vector $\boldsymbol{v}$ in the input space transforms to $T(\boldsymbol{v})$ in the output space, and linearity requires $T(c\boldsymbol{v} + d\boldsymbol{w}) = c\,T(\boldsymbol{v}) + d\,T(\boldsymbol{w})$. Examples: Matrix multiplication $A\boldsymbol{v}$, differentiation in function space.

Linearly dependent $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$. A combination other than all $c_i = 0$ gives $\sum c_i\boldsymbol{v}_i = \boldsymbol{0}$.

Lucas numbers $L_n = 2, 1, 3, 4, \ldots$ satisfy $L_n = L_{n-1} + L_{n-2} = \lambda_1^n + \lambda_2^n$, with eigenvalues $\lambda_1, \lambda_2 = (1 \pm \sqrt{5})/2$ of the Fibonacci matrix $\left[\begin{smallmatrix}1 & 1\\ 1 & 0\end{smallmatrix}\right]$. Compare $L_0 = 2$ with Fibonacci.

Markov matrix M. All $m_{ij} \geq 0$ and each column sum is 1. Largest eigenvalue $\lambda = 1$. If $m_{ij} > 0$, the columns of M^k approach the steady state eigenvector $M\boldsymbol{s} = \boldsymbol{s} > \boldsymbol{0}$.

Matrix multiplication AB. The i, j entry of AB is (row i of A)·(column j of B) = $\sum a_{ik}b_{kj}$. By columns: Column j of $AB = A$ times column j of B. By rows: row i of A multiplies B. Columns times rows: AB = sum of (column k)(row k). All these equivalent definitions come from the rule that AB times $\boldsymbol{x}$ equals A times $B\boldsymbol{x}$.

Minimal polynomial of A. The lowest degree polynomial with $m(A)$ = zero matrix. The roots of m are eigenvalues, and $m(\lambda)$ divides $\det(A - \lambda I)$.

Multiplication $A\boldsymbol{x} = x_1$(column 1) $+ \cdots + x_n$(column n) = combination of columns.

Multiplicities AM **and** GM. The algebraic multiplicity AM of an eigenvalue λ is the number of times λ appears as a root of $\det(A - \lambda I) = 0$. The geometric multiplicity GM is the number of independent eigenvectors (= dimension of the eigenspace for λ).

Multiplier ℓ_{ij}. The pivot row j is multiplied by ℓ_{ij} and subtracted from row i to eliminate the i, j entry: ℓ_{ij} = (entry to eliminate)/(jth pivot).

Network. A directed graph that has constants $c_1, \ldots, c_m$ associated with the edges.

Nilpotent matrix N. Some power of N is the zero matrix, $N^k = 0$. The only eigenvalue is $\lambda = 0$ (repeated n times). Examples: triangular matrices with zero diagonal.

Norm $\|A\|$ **of a matrix**. The "ℓ^2 norm" is the maximum ratio $\|A\boldsymbol{x}\|/\|\boldsymbol{x}\| = \sigma_{\max}$. Then $\|A\boldsymbol{x}\| \le \|A\|\|\boldsymbol{x}\|$ and $\|AB\| \le \|A\|\|B\|$ and $\|A + B\| \le \|A\| + \|B\|$. **Frobenius norm** $\|A\|_F^2 = \sum\sum a_{ij}^2$; ℓ^1 and ℓ^∞ norms are largest column and row sums of $|a_{ij}|$.

Normal equation $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$. Gives the least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$ if A has full rank n. The equation says that (columns of A)·$(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = 0$.

Normal matrix N. $NN^{\mathrm{T}} = N^{\mathrm{T}}N$, leads to orthonormal (complex) eigenvectors.

Nullspace $N(A)$ = Solutions to $A\boldsymbol{x} = \boldsymbol{0}$. Dimension $n - r$ = (# columns) − rank.

Nullspace matrix N. The columns of N are the $n - r$ special solutions to $A\boldsymbol{s} = \boldsymbol{0}$.

Orthogonal matrix Q. Square matrix with orthonormal columns, so $Q^{\mathrm{T}}Q = I$ implies $Q^{\mathrm{T}} = Q^{-1}$. Preserves length and angles, $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$ and $(Q\boldsymbol{x})^{\mathrm{T}}(Q\boldsymbol{y}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$. All $|\lambda| = 1$, with orthogonal eigenvectors. Examples: Rotation, reflection, permutation.

Orthogonal subspaces. Every $\boldsymbol{v}$ in $\boldsymbol{V}$ is orthogonal to every $\boldsymbol{w}$ in $\boldsymbol{W}$.

Orthonormal vectors $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$. Dot products are $\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_j = 0$ if $i \ne j$ and $\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_i = 1$. The matrix Q with these orthonormal columns has $Q^{\mathrm{T}}Q = I$. If $m = n$ then $Q^{\mathrm{T}} = Q^{-1}$ and $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ is an **orthonormal basis** for $\mathbf{R}^n$: every $\boldsymbol{v} = \sum(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{q}_j)\boldsymbol{q}_j$.

Outer product $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ = column times row = rank one matrix.

Partial pivoting. In elimination, the jth pivot is chosen as the largest available entry (in absolute value) in column j. Then all multipliers have $|\ell_{ij}| \le 1$. Roundoff error is controlled (depending on the *condition number* of A).

Particular solution $\boldsymbol{x}_p$. Any solution to $A\boldsymbol{x} = \boldsymbol{b}$; often $\boldsymbol{x}_p$ has free variables = 0.

Pascal matrix P_S = pascal(n). The symmetric matrix with binomial entries $\binom{i+j-2}{i-1}$. $P_S = P_L P_U$ all contain Pascal's triangle with det = 1 (see index for more properties).

Permutation matrix P. There are $n!$ orders of $1, \ldots, n$; the $n!$ P's have the rows of I in those orders. PA puts the rows of A in the same order. P is a product of row exchanges P_{ij}; P is *even* or *odd* (det $P = 1$ or -1) based on the number of exchanges.

Pivot columns of A. Columns that contain pivots after row reduction; not combinations of earlier columns. The pivot columns are a basis for the column space.

Pivot d. The diagonal entry (*first nonzero*) when a row is used in elimination.

Plane (or hyperplane) in $\mathbf{R}^n$. Solutions to $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = 0$ give the plane (dimension $n - 1$) perpendicular to $\boldsymbol{a} \neq \mathbf{0}$.

Polar decomposition $A = QH$. Orthogonal Q, positive (semi)definite H.

Positive definite matrix A. Symmetric matrix with positive eigenvalues and positive pivots. Definition: $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} > 0$ unless $\boldsymbol{x} = \mathbf{0}$.

Projection $\boldsymbol{p} = \boldsymbol{a}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})$ **onto the line through** $\boldsymbol{a}$. $P = \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$ has rank 1.

Projection matrix P **onto subspace** $\boldsymbol{S}$. Projection $\boldsymbol{p} = P\boldsymbol{b}$ is the closest point to $\boldsymbol{b}$ in $\boldsymbol{S}$, error $\boldsymbol{e} = \boldsymbol{b} - P\boldsymbol{b}$ is perpendicular to $\boldsymbol{S}$. $P^2 = P = P^{\mathrm{T}}$, eigenvalues are 1 or 0, eigenvectors are in $\boldsymbol{S}$ or $\boldsymbol{S}^{\perp}$. If columns of A = basis for $\boldsymbol{S}$ then $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$.

Pseudoinverse A^+ **(Moore-Penrose inverse).** The n by m matrix that "inverts" A from column space back to row space, with $\boldsymbol{N}(A^+) = \boldsymbol{N}(A^{\mathrm{T}})$. A^+A and AA^+ are the projection matrices onto the row space and column space. Rank(A^+) = rank(A).

Random matrix rand(n) or randn(n). MATLAB creates a matrix with random entries, uniformly distributed on [0 1] for rand and standard normal distribution for randn.

Rank one matrix $A = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} \neq 0$. Column and row spaces = lines $c\boldsymbol{u}$ and $c\boldsymbol{v}$.

Rank $r(A)$ = number of pivots = dimension of column space = dimension of row space.

Rayleigh quotient $q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}/\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ for symmetric A: $\lambda_{\min} \leq q(\boldsymbol{x}) \leq \lambda_{\max}$. Those extremes are reached at the eigenvectors $\boldsymbol{x}$ for $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$.

Reduced row echelon form R = rref(A). Pivots = 1; zeros above and below pivots; r nonzero rows of R give a basis for the row space of A.

Reflection matrix $Q = I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$. The unit vector $\boldsymbol{u}$ is reflected to $Q\boldsymbol{u} = -\boldsymbol{u}$. All vectors $\boldsymbol{x}$ in the plane mirror $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{x} = 0$ are unchanged because $Q\boldsymbol{x} = \boldsymbol{x}$. The "Householder matrix" has $Q^{\mathrm{T}} = Q^{-1} = Q$.

Right inverse A^+. If A has full row rank m, then $A^+ = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}$ has $AA^+ = I_m$.

Rotation matrix $R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ rotates the plane by θ and $R^{-1} = R^{\mathrm{T}}$ rotates back by $-\theta$. Orthogonal matrix, eigenvalues $e^{i\theta}$ and $e^{-i\theta}$, eigenvectors $(1, \pm i)$.

Row picture of $A\boldsymbol{x} = \boldsymbol{b}$. Each equation gives a plane in $\boldsymbol{R}^n$; planes intersect at $\boldsymbol{x}$.

Row space $\boldsymbol{C}(A^{\mathrm{T}})$ = all combinations of rows of A. Column vectors by convention.

Saddle point of $f(x_1, \ldots, x_n)$. A point where the first derivatives of f are zero and the second derivative matrix ($\partial^2 f/\partial x_i \partial x_j$ = **Hessian matrix**) is indefinite.

Schur complement $S = D - CA^{-1}B$. Appears in block elimination on $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$.

Schwarz inequality $|\boldsymbol{v}\cdot\boldsymbol{w}| \le \|\boldsymbol{v}\|\,\|\boldsymbol{w}\|$.Then $|\boldsymbol{v}^{\mathrm{T}}A\boldsymbol{w}|^2 \le (\boldsymbol{v}^{\mathrm{T}}A\boldsymbol{v})(\boldsymbol{w}^{\mathrm{T}}A\boldsymbol{w})$ if $A = C^{\mathrm{T}}C$.

Semidefinite matrix A. (Positive) semidefinite means symmetric with $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} \ge 0$ for all vectors $\boldsymbol{x}$. Then all eigenvalues $\lambda \ge 0$; no negative pivots.

Similar matrices A **and** B. Every $B = M^{-1}AM$ has the same eigenvalues as A.

Simplex method for linear programming. The minimum cost vector $\boldsymbol{x}^*$ is found by moving from corner to lower cost corner along the edges of the feasible set (where the constraints $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \ge \boldsymbol{0}$ are satisfied). Minimum cost at a corner!

Singular matrix A. A square matrix that has no inverse: $\det(A) = 0$.

Singular Value Decomposition (SVD) $A = U\Sigma V^{\mathrm{T}}$ = **(orthogonal** U**) times (diagonal** Σ**) times (orthogonal** V^{T}**)**. First r columns of U and V are orthonormal bases of $\boldsymbol{C}(A)$ and $\boldsymbol{C}(A^{\mathrm{T}})$ with $A\boldsymbol{v}_i = \sigma_i\boldsymbol{u}_i$ and singular value $\sigma_i > 0$. Last columns of U and V are orthonormal bases of the nullspaces of A^{T} and A.

Skew-symmetric matrix K. The transpose is $-K$, since $K_{ij} = -K_{ji}$. Eigenvalues are pure imaginary, eigenvectors are orthogonal, e^{Kt} is an orthogonal matrix.

Solvable system $A\boldsymbol{x} = \boldsymbol{b}$. The right side $\boldsymbol{b}$ is in the column space of A.

Spanning set $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ **for** $\boldsymbol{V}$. Every vector in $\boldsymbol{V}$ is a combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$.

Special solutions to $A\boldsymbol{s} = \boldsymbol{0}$. One free variable is $s_i = 1$, other free variables $= 0$.

Spectral theorem $A = Q\Lambda Q^{\mathrm{T}}$. Real symmetric A has real λ_i and orthonormal $\boldsymbol{q}_i$ with $A\boldsymbol{q}_i = \lambda_i\boldsymbol{q}_i$. In mechanics the $\boldsymbol{q}_i$ give the *principal axes*.

Spectrum of A = the set of eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$. **Spectral radius** $= |\lambda_{\max}|$.

Standard basis for $\mathbf{R}^n$. Columns of n by n identity matrix (written $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ in $\mathbf{R}^3$).

Stiffness matrix K. If $\boldsymbol{x}$ gives the movements of the nodes in a discrete structure, $K\boldsymbol{x}$ gives the internal forces. Often $K = A^{\mathrm{T}}CA$ where C contains spring constants from Hooke's Law and $A\boldsymbol{x}$ = stretching (strains) from the movements $\boldsymbol{x}$.

Subspace $\boldsymbol{S}$ **of** $\boldsymbol{V}$. Any vector space inside $\boldsymbol{V}$, including $\boldsymbol{V}$ and $\boldsymbol{Z}$ = {zero vector}.

Sum $\boldsymbol{V} + \boldsymbol{W}$ **of subspaces**. Space of all ($\boldsymbol{v}$ in $\boldsymbol{V}$) + ($\boldsymbol{w}$ in $\boldsymbol{W}$). **Direct sum**: $\dim(\boldsymbol{V}+\boldsymbol{W}) = \dim \boldsymbol{V} + \dim \boldsymbol{W}$ when $\boldsymbol{V}$ and $\boldsymbol{W}$ share only the zero vector.

Symmetric factorizations $A = LDL^{\mathrm{T}}$ **and** $A = Q\Lambda Q^{\mathrm{T}}$. The number of positive pivots in D and positive eigenvalues in Λ is the same.

Symmetric matrix A. The transpose is $A^{\mathrm{T}} = A$, and $a_{ij} = a_{ji}$. A^{-1} is also symmetric. All matrices of the form $R^{\mathrm{T}}R$ and LDL^{T} and $Q\Lambda Q^{\mathrm{T}}$ are symmetric. Symmetric matrices have real eigenvalues in Λ and orthonormal eigenvectors in Q.

Toeplitz matrix T. Constant-diagonal matrix, so t_{ij} depends only on $j - i$. Toeplitz matrices represent linear time-invariant filters in signal processing.

Trace of A = sum of diagonal entries = sum of eigenvalues of A. $\operatorname{Tr} AB = \operatorname{Tr} BA$.

Transpose matrix A^{T}. Entries $A^{\mathrm{T}}_{ij} = A_{ji}$. A^{T} is n by m, $A^{\mathrm{T}}A$ is square, symmetric, positive semidefinite. The transposes of AB and A^{-1} are $B^{\mathrm{T}}A^{\mathrm{T}}$ and $(A^{\mathrm{T}})^{-1}$.

Triangle inequality $\|u + v\| \leq \|u\| + \|v\|$. For matrix norms $\|A + B\| \leq \|A\| + \|B\|$.

Tridiagonal matrix T: $t_{ij} = 0$ if $|i - j| > 1$. T^{-1} has rank 1 above and below diagonal.

Unitary matrix $U^{\mathrm{H}} = \overline{U}^{\mathrm{T}} = U^{-1}$. Orthonormal columns (complex analog of Q).

Vandermonde matrix V. $Vc = b$ gives the polynomial $p(x) = c_0 + \cdots + c_{n-1}x^{n-1}$ with $p(x_i) = b_i$ at n points. $V_{ij} = (x_i)^{j-1}$ and $\det V =$ product of $(x_k - x_i)$ for $k > i$.

Vector v in $\mathbf{R}^n$. Sequence of n real numbers $v = (v_1, \ldots, v_n) =$ point in $\mathbf{R}^n$.

Vector addition. $v + w = (v_1 + w_1, \ldots, v_n + w_n) =$ diagonal of parallelogram.

Vector space V. Set of vectors such that all combinations $cv + dw$ remain in V. Eight required rules are given in Section 3.1 for $cv + dw$.

Volume of box. The rows (or columns) of A generate a box with volume $|\det(A)|$.

Wavelets $w_{jk}(t)$ or vectors w_{jk}. Stretch and shift the time axis to create $w_{jk}(t) = w_{00}(2^j t - k)$. Vectors from $w_{00} = (1, 1, -1, -1)$ would be $(1, -1, 0, 0)$ and $(0, 0, 1, -1)$.

MATLAB TEACHING CODES

cofactor	Compute the n by n matrix of cofactors.
cramer	Solve the system $A\boldsymbol{x} = \boldsymbol{b}$ by Cramer's Rule.
deter	Matrix determinant computed from the pivots in $PA = LU$.
eigen2	Eigenvalues, eigenvectors, and $\det(A - \lambda I)$ for 2 by 2 matrices.
eigshow	Graphical demonstration of eigenvalues and singular values.
eigval	Eigenvalues and their multiplicity as roots of $\det(A - \lambda I) = 0$.
eigvec	Compute as many linearly independent eigenvectors as possible.
elim	Reduction of A to row echelon form R by an invertible E.
findpiv	Find a pivot for Gaussian elimination (used by **plu**).
fourbase	Construct bases for all four fundamental subspaces.
grams	Gram-Schmidt orthogonalization of the columns of A.
house	2 by 12 matrix giving the corner coordinates of a house.
inverse	Matrix inverse (if it exists) by Gauss-Jordan elimination.
leftnull	Compute a basis for the left nullspace.
linefit	Plot the least squares fit to m given points by a line.
lsq	Least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$ from $A^{\mathrm{T}}A\hat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$.
normal	Eigenvalues and orthonormal eigenvectors when $A^{\mathrm{T}}A = AA^{\mathrm{T}}$.
nulbasis	Matrix of special solutions to $A\boldsymbol{x} = \boldsymbol{0}$ (basis for nullspace).
orthcomp	Find a basis for the orthogonal complement of a subspace.
partic	Particular solution of $A\boldsymbol{x} = \boldsymbol{b}$, with all free variables zero.
plot2d	Two-dimensional plot for the house figures (cover and Section 7.1).
plu	Rectangular $PA = LU$ factorization with row exchanges.
poly2str	Express a polynomial as a string.
project	Project a vector $\boldsymbol{b}$ onto the column space of A.
projmat	Construct the projection matrix onto the column space of A.
randperm	Construct a random permutation.
rowbasis	Compute a basis for the row space from the pivot rows of R.
samespan	Test whether two matrices have the same column space.
signperm	Determinant of the permutation matrix with rows ordered by p.
slu	LU factorization of a square matrix using *no row exchanges*.
slv	Apply slu to solve the system $A\boldsymbol{x} = \boldsymbol{b}$ allowing no row exchanges.
splu	Square $PA = LU$ factorization *with row exchanges*.
splv	The solution to a square, invertible system $A\boldsymbol{x} = \boldsymbol{b}$.
symmeig	Compute the eigenvalues and eigenvectors of a symmetric matrix.
tridiag	Construct a tridiagonal matrix with constant diagonals a, b, c.

These Teaching Codes are directly available from the Linear Algebra Home Page:
http://web.mit.edu/18.06/www
They were written in MATLAB, and translated into Maple and Mathematica.

LINEAR ALGEBRA IN A NUTSHELL

((A is n by n))

Nonsingular	**Singular**
A is invertible	A is not invertible
The columns are independent	The columns are dependent
The rows are independent	The rows are dependent
The determinant is not zero	The determinant is zero
$A\boldsymbol{x}=\mathbf{0}$ has one solution $\boldsymbol{x}=\mathbf{0}$	$A\boldsymbol{x}=\mathbf{0}$ has infinitely many solutions
$A\boldsymbol{x}=\boldsymbol{b}$ has one solution $\boldsymbol{x}=A^{-1}\boldsymbol{b}$	$A\boldsymbol{x}=\boldsymbol{b}$ has no solution or infinitely many
A has n (nonzero) pivots	A has $r<n$ pivots
A has full rank $r=n$	A has rank $r<n$
The reduced row echelon form is $R=I$	R has at least one zero row
The column space is all of $\mathbf{R}^n$	The column space has dimension $r<n$
The row space is all of $\mathbf{R}^n$	The row space has dimension $r<n$
All eigenvalues are nonzero	Zero is an eigenvalue of A
$A^{\mathrm{T}}A$ is symmetric positive definite	$A^{\mathrm{T}}A$ is only semidefinite
A has n (positive) singular values	A has $r<n$ singular values

Each line of the singular column can be made quantitative using r.