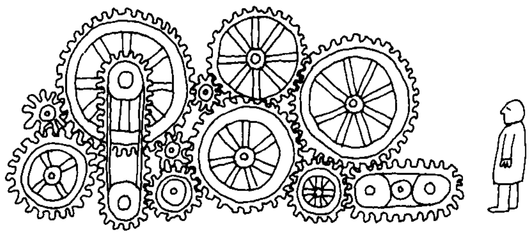
Math 415

Carol Ash



PLUS SOLUTIONS TO ALL PROBLEMS

CHAPTER 1 MATRICES

SECTION 1.1 MATRIX ALGEBRA

matrices

Configurations like

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & \sqrt{5} \end{bmatrix}, \qquad \begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & 2 \\ \pi & 6 & 1/2 \end{bmatrix}, \qquad \begin{bmatrix} 2 \\ 3 \\ -1 \\ 2 \end{bmatrix}$$

are called *matrices*. The numbers inside the matrix are called *entries*.

If the matrix is named B then the entry in row 2, col 3 is usually called b_{23}.

If

$$B = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & -1 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix}$$

then B is 4×3 (4 rows and 3 columns) and $b_{33} = 9$, $b_{21} = 5$, $b_{13} = 4$ etc.

Here's a general $m \times n$ matrix:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The matrices

$$A = \begin{bmatrix} 2 & 5 & 7 \end{bmatrix}$$

$$B = \begin{bmatrix} 4 & -1 & 8 & 9 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

are called *row matrices* or *row vectors*.

The matrices

$$\begin{bmatrix} 4 \\ 5 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

are *column matrices* or *column vectors*.

The matrices

$$P = \begin{bmatrix} 2 & 4 \\ 5 & 6 \end{bmatrix}, \qquad Q = \begin{bmatrix} 5 & 7 & 8 \\ \pi & 2 & -3 \\ 1 & 0 & 9 \end{bmatrix}, \qquad R = \begin{bmatrix} 5 & 6 & 9 & 2 \\ 1 & 2/3 & 0 & e \\ 2 & 1 & 0 & 0 \\ 3 & 3 & 16 & 8 \end{bmatrix}$$

are *square* matrices.

A square matrix has a main diagonal:

$$\begin{bmatrix} \text{main diagonal} \end{bmatrix}$$

For example in matrix Q just above, the diagonal entries are 5,2,9.

A *diagonal* matrix is a square matrix with 0's off the main diagonal, e.g.,

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 4 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$$

The *identity* matrices are

$$[1], \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ etc.}$$

They are all called I.

A square matrix with 0's on one side of the diagonal is called *triangular*. If

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 0 & 0 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 3 & 5 & 9 & 0 \\ 2 & 8 & 3 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 0 & 4 & 5 \\ 0 & 4 & 6 & 9 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

then A and C are *upper triangular* and B is *lower triangular*.

Any matrix with all zero entries is denoted 0 and called a *zero matrix*.

addition and subtraction

Adding and subtracting is done entrywise. If

$$A = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 6 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & \pi & 7 \\ 0 & 2 & 6 \end{bmatrix}$$

then

$$A + B = \begin{bmatrix} 1 & 3+\pi & 12 \\ 1 & 8 & 10 \end{bmatrix} \text{ and } A-B = \begin{bmatrix} 3 & 3-\pi & -2 \\ 1 & 4 & -2 \end{bmatrix}$$

Only same-size matrices can be added and subtracted.

scalar multiplication

To multiply a matrix by a number (i.e., by a scalar) multiply all its entries by the number.

$$\text{If } A = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 6 & 4 \end{bmatrix} \text{ then } 3A = \begin{bmatrix} 6 & 9 & 15 \\ 3 & 18 & 12 \end{bmatrix}$$

$$\text{If } B = \begin{bmatrix} 5 & 6 \\ -1 & 2 \end{bmatrix} \text{ then } -B = \begin{bmatrix} -5 & -6 \\ 1 & -2 \end{bmatrix}$$

properties of addition, subtraction, scalar multiplication

These operations work just the way you would expect.

(1) $A + B = B + A$
(2) $(A + B) + C = A + (B + C)$
(3) $A + 0 = 0 + A = A$
(4) $A - A = 0$
(5) $2(A + B) = 2A + 2B$
(6) $6A + 3A = 9A$
(7) $2(3A) = 6A$

matrix multiplication

The entry in say row 3, col 2 of the product AB is gotten from row 3 of the first factor A and col 2 of the second factor B like this:

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a & b & c \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & p & \cdot & \cdot \\ \cdot & q & \cdot & \cdot \\ \cdot & r & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & ap+bq+cr & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

In other words, dot row 3 of A with col 2 of B to get the row 3, col 2 entry in AB.

If

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \text{ and } B = \begin{bmatrix} 3 & 3 & 4 \\ 1 & 0 & 2 \\ 5 & 10 & 6 \end{bmatrix}$$

then

$$AB = \begin{bmatrix} 1\cdot3+2\cdot1+3\cdot5 & 1\cdot3+2\cdot0+3\cdot10 & 1\cdot4+2\cdot2+3\cdot6 \\ 4\cdot3+5\cdot1+6\cdot5 & 4\cdot3+5\cdot0+6\cdot10 & 4\cdot4+5\cdot4+6\cdot6 \end{bmatrix} = \begin{bmatrix} 20 & 33 & 26 \\ 47 & 72 & 62 \end{bmatrix}$$

Their sizes must be compatible before two matrices can be multiplied. In order to have a product PQ, "across P" must match match "down Q" (i.e., the number of cols in P must be the same as the number of rows of Q).

The order in which two matrices are multiplied makes a difference: the product PQ is not necessarily the same as the product QP.

With the A and B above, AB exists because "across A" matches "down B" but BA doesn't exist because "across B" doesn't match "down A":

$$\textit{can't multiply} \quad \begin{bmatrix} 3 & 3 & 4 \\ 1 & 0 & 2 \\ 5 & 10 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \textit{can't multiply}$$

Here's the formal rule:

Let A be $m \times n$ (m rows and n cols) and B be $p \times q$.
The product AB exists iff $n = p$ in which case AB is $m \times q$.

In other words, if C is $m \times n$ and D is $n \times q$ then CD exists and is $m \times q$

Notation

"iff" means "if and only if"

example 1

If

$$P = \begin{bmatrix} 2 & 3 \\ 1 & 6 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 4 & 3 & 1 \\ 2 & 0 & 2 \end{bmatrix}$$

then

$$PQ = \begin{bmatrix} 14 & 6 & 8 \\ 16 & 3 & 13 \end{bmatrix}$$

and there is no QP.

example 2

If

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 5 & -1 & 2 & 0 \\ 1 & 0 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{bmatrix}$$

then

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

and there is no BA.

example 3

$$[1 \quad 5 \quad 6]\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = [29] \quad (\text{a } 1 \times 1 \text{ matrix})$$

systems of equations in matrix notation

The system of equations

$$\begin{aligned} x + 2y + 3z &= 4 \\ 5x + 6y + 7z &= 8 \\ 9x + 10y + 11z &= 12 \\ 13x + 14y + 15z &= 16 \end{aligned}$$

can be written in matrix form as

$$\begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \\ 13 & 14 & 15 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 12 \\ 16 \end{bmatrix}$$

It can also be written as

$$x\begin{bmatrix} 1 \\ 5 \\ 9 \\ 13 \end{bmatrix} + y\begin{bmatrix} 2 \\ 6 \\ 10 \\ 14 \end{bmatrix} + z\begin{bmatrix} 3 \\ 7 \\ 11 \\ 15 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 12 \\ 16 \end{bmatrix}$$

some observations about matrix multiplication

(1) $$A\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\begin{bmatrix} \text{col 1} \\ \text{of} \\ A \end{bmatrix} + b\begin{bmatrix} \text{col 2} \\ \text{of} \\ A \end{bmatrix} + c\begin{bmatrix} \text{col 3} \\ \text{of} \\ A \end{bmatrix}$$

Look at this 2×2 example to see why it's true. If

$$A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$$

then

$$A\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3a+4b \\ 5a+6b \end{bmatrix} = a\begin{bmatrix} 3 \\ 5 \end{bmatrix} + b\begin{bmatrix} 4 \\ 6 \end{bmatrix} = a\begin{bmatrix} \text{col 1} \\ \text{of} \\ A \end{bmatrix} + b\begin{bmatrix} \text{col 2} \\ \text{of} \\ A \end{bmatrix}$$

(2) $$\text{col 1 of } AB = A\begin{bmatrix} \text{col 1} \\ \text{of} \\ B \end{bmatrix}$$

$$\text{col 2 of } AB = A\begin{bmatrix} \text{col 2} \\ \text{of} \\ B \end{bmatrix} \quad \text{etc.}$$

In other words, if $AB = C$ then

$$A\begin{bmatrix} \text{col 1} \\ \text{of} \\ \text{B} \end{bmatrix} = \begin{bmatrix} \text{col 1} \\ \text{of} \\ \text{C} \end{bmatrix}$$

$$A\begin{bmatrix} \text{col 2} \\ \text{of} \\ \text{B} \end{bmatrix} = \begin{bmatrix} \text{col 2} \\ \text{of} \\ \text{C} \end{bmatrix} \quad \text{etc.}$$

(3) Say M is $n \times 4$ Then

$$M\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \text{col 1} \\ \text{of} \\ \text{M} \end{bmatrix}; \quad M\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \text{col 2} \\ \text{of} \\ \text{M} \end{bmatrix}; \quad M\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \text{col 3} \\ \text{of} \\ \text{M} \end{bmatrix}; \quad M\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \text{col 4} \\ \text{of} \\ \text{M} \end{bmatrix}$$

example 4

If

$$M\begin{bmatrix} 2 & 6 & 7 \\ 3 & 1 & 4 \\ 5 & 8 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 5 & 8 \end{bmatrix}$$

then

$$M\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad M\begin{bmatrix} 6 \\ 1 \\ 8 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \quad M\begin{bmatrix} 7 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$$

some properties of matrix multiplication (assuming that the mult is possible in the first place)

(1) $A(BC) = (AB)C$

In other words, the product ABC can be done by finding AB first or by finding BC first.

Similarly, ABCD can be done in any of the following ways.

Find AB first, then right multiply by C and then by D.
Multiply CD first, then left multiply by B and then by A.
Multiply BC first, then left multiply by A and right multiply by D.
etc.

(2) $A(B + C) = AB + AC$
$(B + C)A = BA + CA$
$(A + B)(C + D) = AC + BC + AD + BD$

(3) $IA = A$, $BI = B$ (where I is an identity matrix)
(4) $(2A)(3B) = 6(AB)$

some non-properties of matrix multiplication

(1) Matrix multiplication is *not* commutative: AB does not necessarily equal BA.

In fact AB and BA might not even both exist and even if they both exist they might not be equal.

Similarly ABC, CBA, BAC, ACB are not necessarily equal.

(2) It is *not* true that if $AB = 0$ then $A = 0$ or $B = 0$.

In other words, it's possible to have $A \neq 0$ and $B \neq 0$ and still have $AB = 0$.

Here's a counterexample.

If $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 \\ 1 & 3 \end{bmatrix}$ then $A \neq 0$ and $B \neq 0$ but $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Here is one consequence of (2). If you want to solve the matrix equation

$$(*) \qquad (X - 3A)(X + 5A) = 0$$

then it is *not* true that $X = 3A$ or $X = -5A$. It *is* true that $X = 3A$ and $X = -5A$ are solutions of (*) but they are not necessarily the *only* solutions because it's possible for $(X-3A)(X+5A)$ to be 0 without either factor itself being 0.

As a special case of (2), it is *not* true that if $B^2 = 0$ then $B = 0$; i.e., it is possible for a nonzero matrix to have a zero square. If $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ then $B \ne 0$ but $B^2 = 0$.

(3) It is *not* true that if $AB = AC$ and $A \ne 0$ then $B = C$; i.e., you can't always cancel matrices. As a counterexample,

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 3 & 0 \\ 6 & 0 \end{bmatrix}}_{C}$$

(because both sides come out to be $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$) but you can't cancel the A's and conclude that $B = C$.

warning

1. $3A^2 + 2A$ factors into $A(3A + 2I)$, *not* into $A(3A + 2)$. The latter notation is wrong since you can't add the matrix $3A$ and the number 2.
2. $(A + B)^2$ expands to $A^2 + AB + BA + B^2$, *not* to $A^2 + 2AB + B^2$.
3. $CAD + CBD$ factors into $C(A + B)D$, *not* into $CD(A + B)$
4. $(B+C)A$ is $BA + CA$ *not* $AB + AC$.
4. Start with a matrix A.
Left multiplying it by B means computing BA.
Right multiplying it by B means computing AB.
It's ambiguous to talk about "multiplying by B" because BA and AB are not necessarily equal.

PROBLEMS FOR SECTION 1.1 (solutions at the end of the notes)

1. Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 3 & 4 \\ -5 & 6 & 7 \end{bmatrix}$.
Find AB, BA, A + B, 2A (if they exist).

2. Let

$$A = \begin{bmatrix} 3 & 4 \\ 3 & 6 \\ 1 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \quad Q = \begin{bmatrix} 5 & 0 & 2 \\ 3 & 4 & 1 \\ 2 & 2 & 2 \end{bmatrix}$$

$$F = [2 \quad 1 \quad 3], \quad H = \begin{bmatrix} 1 & 1 & 5 \\ 2 & 3 & 2 \end{bmatrix}$$

Which products exist.

3. Suppose A is 3×7 and B is 7×12. What size must C be if
(a) ABC is to exist
(b) CAB is to exist

4. Find AB and BA (if they exist) where $A = \begin{bmatrix} 2 & 3 & 6 \\ 1 & 7 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 2 & 6 \\ 1 & 1 & 1 \end{bmatrix}$.

5. Write the system of equations in matrix form:

$$\begin{aligned} 2x + 3y + 5z &= 7 \\ x - y + 6z &= 8 \end{aligned}$$

6. Given a 6×3 matrix A and a 3×7 matrix B. What does $\sum_{j=1}^{3} a_{2j}\, b_{j5}$ compute.

7. Let $\vec{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and let M be 3×3. What is $M\vec{i}$.

8. Find AB and BA if possible.

(a) $A = [2 \quad 3 \quad 4]$ and $B = \begin{bmatrix} 5 \\ 6 \\ 1 \end{bmatrix}$

(b) $A = [2 \quad 3]$ and $B = \begin{bmatrix} 5 \\ 6 \\ 1 \end{bmatrix}$

9. Suppose

$$B \begin{bmatrix} 1 & 2 & 5 \\ 3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 9 & 8 \\ 3 & 1 & 6 \\ 4 & 2 & 7 \end{bmatrix}.$$

(a) What size is the matrix B.

(b) Find $B \begin{bmatrix} 5 \\ 2 \end{bmatrix}$.

10. If A and B are same size square matrices then AB and BA both exist. Can AB and BA both exist if A and B are not square?

11. If $AB = 5B$ find $A^4 B$.

12. Show that any two 3×3 diagonal matrices commute; i.e., if A and B are diagonal then $AB = BA$.

13. Let $D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$.

Describe the effect of
(a) left multiplying an arbitrary matrix by D
(b) right multiplying an arbitrary matrix by D.

14. Suppose A and B are 4×4 matrices and $Ax = Bx$ for all 4×1 column matrices x. Show that $A = B$.

Suggestion: Think about the col matrices $i = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $j = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ etc.

15. Show that if A and B commute (i.e., $AB = BA$) then the two matrices $2A + 3B$ and $6A - 4B$ also commute.

16. Does $(A + B)(A - B)$ equal $A^2 - B^2$.

17. Factor (a) $A + AB$ (b) $A + BA$

18. A matrix Q is a square root of B if $Q^2 = B$.

(a) To find square roots of I consider solving $Q^2 = I$ as follows.

$Q^2 = I$
$Q^2 - I = 0$
$(Q - I)(Q + I) = 0$
$Q - I = 0$ or $Q + I = 0$
$Q = I$ or $Q = -I$
So there are two square roots of I, namely I and $-I$.

Find the mistake in this reasoning.

(b) The matrix I has lots of square roots.
Show that all the following are square roots of I.

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 7 \\ 1/7 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 8 \\ 1/8 & 0 \end{bmatrix}$$

19. Factor if possible.
(a) $AB + BC$ (b) $AB + CB$ (c) $A^2B + AB^2$ (d) $A^3 + 2A^2 + 6A$
(e) $A^2 + 3A + 2I$ (f) $A^2 + 3AB + 2B^2$

20. Suppose $A \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = C$.

Write an equation relating column 1 of C with the columns of A.

SECTION 1.2 ROW AND COLUMN OPERATIONS

row operations

Here's a list of row operations which will be used at various times during the term to get a new and, we hope, simpler matrix which in turn can be used to answer questions about the original matrix. (There's a similar list of column operations.)

(1) Interchange two rows.
(2) Multiply a row by a nonzero constant.
(3) Add a multiple of one row to another row.

For example, let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 9 \\ 0 & 2 & 4 \end{bmatrix}$$

The row operation "add twice row 1 to row 2", abbreviated $R_2 = 2R_1 + R_2$ produces the new matrix

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 12 & 15 \\ 0 & 2 & 4 \end{bmatrix}$$

row equivalent matrices

Every row operation has an "opposite" operation.

For instance in the example above, I got B from A above by doing the row op $R_2 = 2R_1 + R_2$ and you can get A from B by doing the opposite row op $R_2 = -2R_1 + R_2$.

The opposite of the row operation $R_4 = 5R_4$ is $R_4 = \frac{1}{5} R_4$.

The opposite of the row operation "switch rows 2 and 3" is "switch rows 2 and 3".

If you get matrix P from matrix Q by doing row operations then you can also get Q from P by doing the opposite row operations; in that case the two matrices are called *row equivalent* and we write $P \sim Q$.

reduced echelon form (also called row echelon form)

Every matrix can be row operated into a simple form called reduced echelon form, consisting for the most part of columns like these (in this order)

$$\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix} \qquad \text{etc.}$$

The 1's are called *pivots*. Other columns may interrupt the cols with pivots, but only in a particular way:

After the column $\begin{matrix} 0 \\ 1 \\ 0 \\ 0 \end{matrix}$ you can have columns of the form $\begin{matrix} a \\ b \\ 0 \\ 0 \end{matrix}$.

After the column $\begin{matrix} 0 \\ 0 \\ 1 \\ 0 \end{matrix}$ you can have columns of the form $\begin{matrix} a \\ b \\ c \\ 0 \end{matrix}$ etc.

Here are three examples of reduced echelon form, with the pivots boxed.

$$\begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \boxed{1} & 0 & 2 & 0 & -2 & -1 & 0 \\ 0 & \boxed{1} & 1 & 0 & 6 & 1 & 0 \\ 0 & 0 & 0 & \boxed{1} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & \boxed{1} & 8 & 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & \boxed{1} & 0 & 3 & 8 \\ 0 & 0 & 0 & 0 & \boxed{1} & 4 & 9 \end{bmatrix}$$

Here are some examples of matrices *not in reduced echelon form*

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{(col 3 is no good)}$$

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{(col 2 is no good)}$$

> The best way to use row ops to reach echelon form is to work on the entries in col 1 first, then col 2, then col 3 etc. Within each col, try to get the 1 (the pivot) first and then use the pivot to get 0's above and below

example 1

Let

$$A = \begin{bmatrix} 2 & 4 & -2 & 8 & 4 \\ 3 & 6 & 1 & 12 & -2 \\ 9 & 18 & 1 & 36 & 38 \end{bmatrix}$$

To find the echelon form of A, try to get a 1 in the 1,1 spot (i.e., in row 1, col 1) and 0's under it:

$$A \sim \begin{bmatrix} 1 & 2 & -1 & 4 & 2 \\ 3 & 6 & 1 & 12 & -2 \\ 9 & 18 & 1 & 36 & 38 \end{bmatrix} \qquad R_1 = \tfrac{1}{2} R_1$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 4 & 2 \\ 0 & 0 & 4 & 0 & -8 \\ 0 & 0 & 10 & 0 & 20 \end{bmatrix} \qquad \begin{aligned} R_2 &= -3R_1 + R_2 \\ R_3 &= -9R_1 + R_3 \end{aligned}$$

Now try to get a 1 in the 2,2 spot. But the 0 already there can never turn into a 1 by a row op without undoing the first col. So give up on col 2 and try to get a 1 in the 2,3 spot with 0's above and below:

$$A \sim \begin{bmatrix} 1 & 2 & -1 & 4 & 2 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 10 & 0 & 20 \end{bmatrix} \qquad R_2 = \tfrac{1}{4} R_2$$

$$\sim \begin{bmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 40 \end{bmatrix} \qquad \begin{aligned} R_1 &= R_{12} + R_1 \\ R_3 &= -10R_2 + R_3 \end{aligned}$$

It's impossible to get a 1 in the 3,4 spot (row 3, col 4) without undoing earlier achievements so try for a 1 in the 3,5 spot with 0's above:

$$A \sim \underbrace{\begin{bmatrix} \boxed{1} & 2 & 0 & 4 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}}_{\text{reduced echelon form}} \qquad \begin{aligned} R_3 &= \tfrac{1}{40} R_3 \\ R_2 &= 2R_3 + R_2 \end{aligned}$$

unreduced echelon form

Reduced form requires pivots of 1's with 0's above and below. An *un*reduced form allows non-zero pivots and requires 0's below the pivots only. Here's an example of an unreduced echelon form:

$$\begin{matrix} \boxed{2} & 3 & 5 & 7 & 1 & 5 \\ 0 & 0 & \boxed{6} & 8 & 2 & 8 \\ 0 & 0 & 0 & 0 & \boxed{3} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

If

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

then the row ops

$$R_2 = -2R_1 + R_2, \quad R_3 = R_1 + R_3, \quad R_3 = 3R_2 + R_3$$

produce the *un*reduced echelon form

$$\begin{bmatrix} \boxed{2} & 1 & 1 & 1 \\ 0 & \boxed{-1} & -2 & -4 \\ 0 & 0 & \boxed{-4} & -4 \end{bmatrix}$$

Every matrix has many unreduced echelon forms but it can be shown that *there is a unique reduced echelon form*. If you continue row operating on A you'll get the unique reduced form

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

warning

1. Performing the row op $R_2 = 2R_1 + R_2$ changes R_2 (it addes twice row 1 to row 2) but it doesn't change R_1. It does *not* double R_1.
2. Don't use the incomplete notation $6R_1 + R_2$ if you really mean $R_2 = 6R_1 + R_2$.
3. If A row operates to B, don't write A = B. Write A ~ B.

elementary matrices

If a row or col operation (just *one*) is performed on I, the result is called an elementary matrix.

For example, if you do the row operation $R_1 \leftrightarrow R_2$ on $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ you get the elementary matrix

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

If you do the column operation $C_2 = 2\,C_3 + C_2$ on the 3×3 matrix I, you get the elementary matrix

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

There's a connection between doing row and col operations on a matrix M and multiplying M by elementary matrices (assume all matrices are same size square):

> If an elementary matrix E is created from I by doing a *row* operation then *left* multiplying M by E will have the effect of doing that same row operation on M.
> If an elementary matrix E is created from I by doing a *col* operation then *right* multiplying M by E will have the effect of doing that same col operation on M.

Look at an example to see the idea. Let

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Use E_1 and E_2 from above. Then

$$E_1 M = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix}$$

Left multiplying by E_1 did the row op $R_1 \leftrightarrow R_2$ to M, the row op that created E_1 in the first place. Similarly,

$$M E_2 = \begin{bmatrix} a & b+2c & c \\ d & e+f & f \\ g & h+2i & i \end{bmatrix}$$

Right multiplying by E_2 did the col op $C_2 = 2\,C_3 + C_2$ to M, the same col op that created E_2.

example 2

Suppose A is 3×3 and the operations $C_1 \leftrightarrow C_2$, $R_2 = 2R_1 + R_2$, $R_2 \leftrightarrow R_3$ are performed successively to get a new matrix B. Then

$$B = E_3\, E_2\, A\, E_1$$

where

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

PROBLEMS FOR SECTION 1.2

1. Which of the following are (reduced) echelon form.

(a) $\begin{matrix} 1 & 3 & 0 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}$ (b) $\begin{matrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix}$ (c) $\begin{matrix} 1 & 3 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix}$

(d) $\begin{matrix} 1 & 0 & 0 & 5 & 8 \\ 0 & 1 & 0 & 6 & 9 \\ 0 & 0 & 1 & 7 & 2 \end{matrix}$ (e) $\begin{matrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix}$ (f) $\begin{matrix} 0 & 1 & 2 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{matrix}$

2. Use row operations to find the reduced echelon form for each matrix.

(a) $\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & 3 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 6 & 5 \\ 4 & 10 & 0 \\ 1 & 10 & 0 \end{bmatrix}$

3. Are these legitimate row operations?
(a) replace row 3 by row 3 + 2 row 4
(b) replace row 3 by row 4 + 2 row 3
(c) replace row 3 by row 1 + 2 row 5

4. If the row ops $R_2 = 2R_1 + R_2$, $R_2 = 5R_2$, $R_1 \leftrightarrow R_2$ turn A into B, what row ops (and in what order) take B back to A.

5. Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Show that A is an elementary matrix and then find A^{17} easily without doing any actual multiplying.

6. Start with a 2×2 matrix A and do the row op $R_2 = 3R_1 + R_2$. Call the resulting matrix B.
(a) Then B = EA for what matrix E.
(b) And A = FB for what matrix F.

7. Let A be the matrix in #2(b).
(a) Find B so that BA is the echelon form of A.
(b) Is the matrix B from part (a) an elementary matrix.

8. Find the following product easily, without actually multiplying, by thinking about elementary matrices.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ x & x & x & x \\ y & y & y & y \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{bmatrix}$$

9. Are these unreduced echelon forms. If so, identify the cols with pivots.

(a) $\begin{matrix} 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}$ (b) $\begin{matrix} 2 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$ (c) $\begin{matrix} 2 & 0 & 3 \\ 0 & 0 & 4 \\ 0 & 1 & 5 \end{matrix}$

(d) $\begin{array}{ccccc} 5 & 2 & 4 & 7 & 9 \\ 0 & 3 & 6 & 8 & 10 \\ 0 & 0 & 0 & 0 & 11 \\ 0 & 0 & 0 & 0 & 12 \end{array}$ (e) $\begin{array}{ccccc} 5 & 2 & 4 & 7 & 9 \\ 0 & 3 & 6 & 8 & 10 \\ 0 & 0 & 0 & 0 & 11 \\ 0 & 0 & 0 & 0 & 0 \end{array}$

SECTION 1.3 DETERMINANTS

This whole section is about *square* matrices.

definition of the determinant

For every square matrix M there is a number called the determinant of M, denoted det M or $|M|$. I'll illustrate the definition for the 5×5 matrix in Figs 1 and 2 below.

First form products of five factors by choosing one entry from each column without repeating any rows. One such product (Fig 1) is $11 \cdot 7 \cdot 18 \cdot 70 \cdot 5$. Another product (Fig 2) is $6 \cdot 2 \cdot 13 \cdot 70 \cdot 20$.

For a 5×5 matrix there are 5! such products.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & \boxed{5} \\ 6 & \boxed{7} & 8 & 9 & 10 \\ \boxed{11} & 12 & 13 & 14 & 15 \\ 16 & 17 & \boxed{18} & 19 & 20 \\ 40 & 50 & 60 & \boxed{70} & 80 \end{bmatrix} \qquad \begin{bmatrix} \boxed{1} & 2 & 3 & 4 & 5 \\ 6 & \boxed{7} & 8 & 9 & 10 \\ 11 & 12 & 13 & \boxed{14} & 15 \\ 16 & 17 & 18 & 19 & \boxed{20} \\ 40 & 50 & \boxed{60} & 70 & 80 \end{bmatrix}$$

FIG 1 FIG 2

Next, give each product a sign as follows.

Write down the row number of each factor, reading left to right. This will be a permutation (rearrangement) of the numbers 1,2,3,4,5.

Identify each pair of numbers in the permutation that goes "down". This is called an *inversion*.

For instance, for the product selected in Fig 1 the rows numbers are 3 2 4 5 1. Fig 3 shows that there are 5 inversions in the permutation 3 2 4 5 1.

For instance, the permutation of row numbers corresponding to the product selected in Fig 2 is 1 2 5 3 4 and Fig 4 shows 2 inversions.

If the number of inversions in a permutation is even then the product is given a plus sign. If the number of inversions is odd then the product gets a minus sign.

Finally, add up all the 5! signed products. The sum is det M.

3 2 4 5 1

5 inversions
FIG 3

1 2 5 3 4

2 inversions
FIG 4

example 1

Let

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

The table below shows all 6 products and the number of inversions for each corresponding permutation of row numbers.

product	row numbers	number of inversions	sign of product
aei	123	0	+
ahf	132	1	−
dbi	213	1	−
dhc	231	2	+
gbf	312	2	+
gec	321	3	−

So $|A| = aei - ahf - dbi + dhc + gbf - gec$

the determinant of a 1 × 1 matrix

If $M = [a]$ then $\det M = a$ (there's only one product in the sum and it gets a plus sign).

the determinant of a 2 × 2 matrix

If $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $|M| = ad - bc$.

proof

The two possible products are ad and bc.
The product ad has row numbers 12. There are 0 inversions so the attached sign is +.
The product bc has row numbers 21. There is 1 inversion so the attached sign is −.
So $\det M = ad - bc$.

the cofactor of an entry

Every entry in a square matrix has a *minor* and a *cofactor*, defined as follows.

The minor of an entry is the determinant you get by deleting the row and col of that entry .

The cofactor of an entry is its minor with a sign attached according to the entry's location in the checkerboard pattern in Fig 5. (More generally, the cofactor of the entry in row i, col j has sign $(-1)^{i+j}$.)

$$\begin{matrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{matrix}$$

FIG 5

For example if

$$A = \begin{bmatrix} 1 & 2 & 7 & 8 \\ 4 & 3 & 1 & 9 \\ 5 & 2 & 9 & 6 \\ 8 & 1 & 3 & 8 \end{bmatrix}$$

then to get the minor of the 7 in row 1, col 3, form a determinant by deleting row 1 and col 3 :

$$\text{minor of the 7 in row 1, col 3} = \begin{vmatrix} 1 & 2 & 7 & 8 \\ 4 & 3 & 1 & 9 \\ 5 & 2 & 9 & 6 \\ 8 & 1 & 3 & 8 \end{vmatrix} = \begin{vmatrix} 4 & 3 & 9 \\ 5 & 2 & 6 \\ 8 & 1 & 8 \end{vmatrix}$$

To get the cofactor, choose the sign in row 1 col 3 of the checkerboard in Fig 5.

$$\text{cofactor of the 7 in row 1, col 3} = +\begin{vmatrix} 4 & 3 & 9 \\ 5 & 2 & 6 \\ 8 & 1 & 8 \end{vmatrix}$$

(In this case the minor and the cofactor are the same.) Similarly,

$$\text{minor of the 1 in row 2, col 3} = \begin{vmatrix} 1 & 2 & 8 \\ 5 & 2 & 6 \\ 8 & 1 & 8 \end{vmatrix}$$

$$\text{cofactor of the 1 in row 2, col 3} = -\begin{vmatrix} 1 & 2 & 8 \\ 5 & 2 & 6 \\ 8 & 1 & 8 \end{vmatrix}$$

computing the determinant by expansion by minors

To find the det of say a 3×3 matrix M (same idea for an $n \times n$), pick any row. Then

$|M|$ = 1st entry in the row $\times$ its cofactor
+ 2nd entry in the row $\times$ its cofactor
+ 3rd entry in the row $\times$ its cofactor

You can also find the det by picking any col. Then

$|M|$ = 1st entry in the col $\times$ its cofactor
+ 2nd entry in the col $\times$ its cofactor
+ 3rd entry in the col $\times$ its cofactor

Here's a determinant found by expanding down column 2:

$$\begin{vmatrix} 10 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = -2\underbrace{\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix}}_{-6} + 5\underbrace{\begin{vmatrix} 10 & 3 \\ 7 & 9 \end{vmatrix}}_{69} - 8\underbrace{\begin{vmatrix} 10 & 3 \\ 4 & 6 \end{vmatrix}}_{48}$$

$$= -27$$

I'm leaving out the proof that expansion by minors agrees with the method given in the definition of a determinant (it's very tedious).

example 2

Let

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

Here's an expansion down column 4, a good column to use because it has some zero entries:

$$\det A = 2\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 2 & 1 & 4 \end{vmatrix} - 3\begin{vmatrix} 1 & 2 & 3 \\ 1 & -1 & 3 \\ 2 & 1 & 4 \end{vmatrix}$$

Now work on each 3×3 det. Expanding the first one down col 1 and the second across row 1 we get

$$\det A = 2\left(1\begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} + 2\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix}\right) - 3\left(1\begin{vmatrix} -1 & 3 \\ 1 & 4 \end{vmatrix} - 2\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} + 3\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix}\right)$$

$$= 2(2 + 2) - 3(-7 + 4 + 9)$$
$$= -10$$

some properties of determinants

(1) The determinant of a product is the product of the determinants, i.e.,

$$|AB| = |A||B|$$

For example, if $|A| = 2$ and $|B| = 3$ then $|AB| = 6$.

As a special case,

$$|A^n| = |A|^n;$$

e.g., cubing a matrix will cube its determinant.

(2) If A is $n \times n$ and c is any number then

$$|cA| = c^n |A|$$

This says that if A is 7×7 and every entry of A is quadrupled, then the determinant of A is multiplied by 4^7.

proof

The proof of (1) takes too long.

Here's a proof of (2) in the 3×3 case. Let

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

By the definition of determinants, det A is a sum of signed products.
Back in example 1 there's a table of all the signed products, and we got

$$|A| = aei - ahf - dbi + dhc + gbf - gec$$

When you find the determinant of say $4A = \begin{bmatrix} 4a & 4b & 4c \\ 4d & 4e & 4f \\ 4g & 4h & 4i \end{bmatrix}$ you get the same table, but each product has three extra 4's in it:

	row numbers	number of inversions	sign of product
$4a \cdot 4e \cdot 4i$	124	0	+
$4a \cdot 4h \cdot 4f$	132	1	−
$4d \cdot 4b \cdot 4i$	213	1	−
$4d \cdot 4h \cdot 4c$	231	2	+
$4g \cdot 4b \cdot 4f$	312	2	+
$4g \cdot 4e \cdot 4c$	321	3	−

So $|4A| = 4a \cdot 4e \cdot 4i - 4a \cdot 4h \cdot 4f - 4d \cdot 4b \cdot 4i + 4d \cdot 4h \cdot 4c + 4g \cdot 4b \cdot 4f - 4g \cdot 4e \cdot 4c$

$$= 4^3 (aei - ahf - dbi + dhc + gbf - gec)$$
$$= 4^3 |A|. \quad \text{QED}$$

warning

1. It is not a rule that $|A + B| = |A| + |B|$ (see problem 5)
2. It is not a rule that $|3A| = 3|A|$. Rather, if A is $n \times n$ then $|3A| = 3^n |A|$.
3. The rule is not that $|-A| = |A|$, nor is the rule $|-A| = -|A|$.
   Rather, if A is $n \times n$ then $|-A| = (-1)^n |A|$.

determinant of diagonal and triangular matrices

The determinant of a diagonal or a triangular matrix is the product of the diagonal entries.

As a special case, $|I| = 1$.

For example,

$$\begin{vmatrix} 2 & 0 & 0 \\ 3 & 4 & 0 \\ 50 & 6 & 7 \end{vmatrix} = 2\cdot 4\cdot 7 = 56$$

To see why this works, just expand across row 1. Or use the original definition of determinant: there is only one nonzero product, namely $2\cdot 4\cdot 7$ and its sign is +.

the effect of row (and col) operations on a determinant

(1) Multiplying a row (or col) by a number will multiply the entire det by that number.

If say row 2 is doubled then new det = 2 × old det; e.g.,

$$\begin{vmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

From another point of view, this says that a common factor can be pulled out of a row or a col.

For example,

$$\begin{vmatrix} 6 & 3 & 9 \\ 2 & 1 & 5 \\ 2 & 3 & 1 \end{vmatrix} = 2\begin{vmatrix} 3 & 3 & 9 \\ 1 & 1 & 5 \\ 1 & 3 & 1 \end{vmatrix} \quad \text{pull 2 out of col 1}$$

$$= 6\begin{vmatrix} 1 & 1 & 3 \\ 1 & 1 & 5 \\ 1 & 3 & 1 \end{vmatrix} \quad \text{pull 3 out of row 1}$$

(2) Adding a multiple of one row to another row (or a multiple of one col to another col) doesn't change the determinant.

(3) Interchanging two rows (or two cols) changes the sign of the determinant.

For example,

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 5 \\ 7 & 7 & 7 \end{vmatrix} = -\begin{vmatrix} 7 & 7 & 7 \\ 4 & 5 & 5 \\ 1 & 2 & 3 \end{vmatrix} \quad \text{(rows 1 and 3 were switched)}$$

(4) Row and col operations preserve the zero-ness and nonzero-ness of a determinant. In other words, if a det is 0 then after row ops the det is still 0; if a det is nonzero then after row ops it's still nonzero.

Row and col operations can multiply the determinant by a nonzero scalar, and in particular can change the sign of the determinant, but property (4) says that they can't change a zero det to nonzero or vice versa.

proof of (1)

I'll compare $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$ with $\begin{vmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{vmatrix}$.

By the definition of a det, the second determinant is the sum of terms

(*) $\quad a\cdot h\cdot 2f$
$\quad 2d\cdot h\cdot c$
$\quad$ etc

with appropriate signs.

The first determinant is the sum of terms

$$(**)\qquad \begin{array}{l} a\cdot h\cdot f \\ d\cdot h\cdot c \\ \text{etc.} \end{array}$$

with the same signs respectively as in (*).
Each term in (*) is twice the corresponding term in (**). So the second determinant is twice the first det. QED

proof of (2)
I'll illustrate the idea in the 3×3 case for one particular row operation.
Let A be 3×3.
Let B be the matrix obtained from A by doing the row operation $R_3 = 2R_1 + R_3$.
Then $B = EA$ where $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$, the elementary matrix obtained from I by doing the row operation. By property (1) of determinants,

$$|B| = |EA| = |E|\,|A|$$

But E is a diagonal matrix and $|E| = 1\cdot 1\cdot 1 = 1$. So $|B| = |A|$. QED

proof of (3)
I'll illustrate the idea in the 3×3 case for one particular row operation.
Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. I'll interchange rows 1 and 3 and find the new determinant.

The interchange can be accomplished in a round-about way by doing the following four row operations:

$$R_1 = -R_3 + R_1 \qquad \begin{array}{ccc} a-g & b-h & c-i \\ d & e & f \\ g & h & i \end{array}$$

$$R_3 = -R_1 + R_3 \qquad \begin{array}{ccc} a-g & b-h & c-i \\ d & e & f \\ a & b & c \end{array}$$

$$R_1 = -R_3 + R_1 \qquad \begin{array}{ccc} -g & -h & -i \\ d & e & f \\ a & b & c \end{array}$$

$$R_1 = -1\cdot R_1 \qquad \begin{array}{ccc} g & h & i \\ d & e & f \\ a & b & c \end{array}$$

As already proved, the first three row ops leave the determinant unchanged and the last row op multiplies the determinant by -1. So all in all, interchanging rows 1 and 3 changed the sign of the determinant. QED

example 3
You can use row and col operations to get 0 entries in a matrix, so that there's less work involved in the expansion by minors.
For instance,

$$\begin{vmatrix} 198 & 0 & 99 & 99 \\ 1 & 1 & -2 & 0 \\ -1 & 0 & 5 & 2 \\ 1 & -3 & 6 & 1 \end{vmatrix} = 99 \begin{vmatrix} 2 & 0 & 1 & 1 \\ 1 & 1 & -2 & -0 \\ -1 & 0 & 5 & 2 \\ 1 & -3 & 6 & 1 \end{vmatrix} \quad \text{(pull 99 out of row 1)}$$

$$= 99 \begin{vmatrix} 2 & 0 & 1 & 1 \\ 1 & 1 & -2 & 0 \\ -1 & 0 & 5 & 2 \\ 4 & 0 & 0 & 1 \end{vmatrix} \quad \text{(do the row op } R_4 = 3R_2 + R_4\text{)}$$

$$= 99 \begin{vmatrix} 2 & 1 & 1 \\ -1 & 5 & 2 \\ 4 & 0 & 1 \end{vmatrix} \quad \text{(expand down col 2)}$$

$$= 99\left(4 \begin{vmatrix} 1 & 1 \\ 5 & 2 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ -1 & 5 \end{vmatrix} \right) \text{(expand across row 3)}$$

$$= -99$$

warning

Suppose $M = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$.

If you do the row op "add −2 row 1 to row 2" you get the new matrix $\begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix}$.

The new matrix does not equal the old one.So *don't* write $\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix}$.
If you have to write something, use a "row equivalent" symbol and write
$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \sim \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix}$

But the new *determinant* does equal the old *determninant* so you *can* write

$$\begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 0 & -1 \end{vmatrix}.$$

computing a det by triangularizing first

A standard way to compute a determinant efficiently is to use row ops to convert to triangular form.
Let

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

The row ops

$$R_2 = -2R_1 + R_2$$
$$R_3 = -3R_1 + R_3$$
$$R_3 = -\frac{5}{3} R_2 + R_3$$

change A to triangular form and they're the type of row ops that don't change the value of the determinant so

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -5 \\ 0 & 0 & -\frac{2}{3} \end{vmatrix} = 1 \cdot -3 \cdot -\frac{2}{3} = 2 \quad \text{(the product of the diagonal entries)}.$$

incomplete list of ways to spot zero determinants

(1) If there is an all zero row or all zero col then the det is 0.
(2) If two rows are identical or two cols are identical then the det is 0.
(3) If one row is a multiple of another row or one col is a multiple of another col then the det is 0.
(But remember, these are not the *only* ways in which a det can be 0.)
For example,

$$\begin{vmatrix} 6 & 9 & 12 \\ 4 & 6 & 8 \\ 7 & 2 & \pi \end{vmatrix} = 0$$

because row 2 is a multiple of row 1; in particular row 2 = $\frac{2}{3}$ row 1).

proof of (1)

The det is 0 because you can expand by minors across the all zero row or col.

proof of (2)

If say R_2 and R_3 are identical, you can add $-R_2$ to R_3. This doesn't change the det but now row 3 is all 0's So the det is 0 by (1).

proof of (3)

If say row 2 = 7 row 4 you can do the row op row 2 = −7 row 4 + row 2. This doesn't change the det but now row 2 is all 0's. So the det is 0 by (1).

warning

Use the right notation to distinguish a matrix from its determinant. The notations

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and even $\begin{matrix} a & b \\ c & d \end{matrix}$ (when I get lazy) mean matrix. The notation $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ means det.

Letters such as A and M are often used for matrices in which case their determinants are $|A|$, $|M|$ or det A, det M.

PROBLEMS FOR SECTION 1.3

1. Look at the matrix back in Fig 1. According to the definition, the determinant is a sum of signed products. What sign do the following products get
   (a) $1 \cdot 12 \cdot 60 \cdot 19 \cdot 10$ (b) $40 \cdot 17 \cdot 13 \cdot 9 \cdot 5$

2. (a) Find $\begin{vmatrix} 2 & -4 \\ 6 & 3 \end{vmatrix}$.

   (b) Find $\begin{vmatrix} 2 & 0 & 3 \\ 10 & 1 & 17 \\ 7 & 12 & -4 \end{vmatrix}$ by expanding across row 3.

   (c) Find $\begin{vmatrix} 2 & 0 & 1 & 1 \\ 1 & 1 & -2 & 0 \\ -1 & 0 & 5 & 2 \\ 1 & -3 & 6 & 1 \end{vmatrix}$ by expanding across row 4.

   (d) Find $\begin{vmatrix} 2 & 3 & 4 & 7 \\ 0 & 3 & 1 & 6 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 5 \end{vmatrix}$.

3. Look at

$$\begin{vmatrix} 12 & 3 & 4 & \boxed{d} \\ \boxed{a} & 6 & -3 & 21 \\ 8 & \boxed{b} & 1 & 3 \\ 4 & 1 & \boxed{c} & 6 \end{vmatrix}$$

The det is a sum of signed products. One of the products is abcd. Find its attached sign

(a) using the definition of $|A|$

(b) by expanding down col 2 and keeping track of how the product $5 \cdot 9 \cdot 2 \cdot 7$ turns up

4. If you want more practice computing determinants let

$$A = \begin{bmatrix} 3 & 1 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}$$

and find $|A|$, $|B|$, AB, $|AB|$ to see that $|AB|$ really equals $|A||B|$.

5. Find a specific counterexample to disprove the following:

$|A + B| = |A| + |B|$ for all (same size square) A,B.

6. Suppose M is 37×37. What sign is in the row 27, col 29 spot on the checkerboard pattern in Fig 5.

7. Suppose A and B are 4×4, $|A| = 6$, $|B| = -\frac{1}{2}$. Find (if possible)

(a) $|AB|$ (b) $|A| + |B|$ (c) $|A+B|$ (d) $|3A|$ (e) $3|A|$ (f) $\det \frac{A}{|A|}$

8. Find $|-A|$ if $|A| = -7$ and (a) A is 4×4 (b) A is 5×5

9. (a) Does $\begin{vmatrix} 0 & 0 & 5 \\ 0 & 6 & 0 \\ 7 & 0 & 0 \end{vmatrix}$ equal $5 \cdot 6 \cdot 7$.

(b) Does $\begin{vmatrix} 0 & 0 & 0 & 5 \\ 0 & 0 & 6 & 0 \\ 0 & 7 & 0 & 0 \\ 8 & 0 & 0 & 0 \end{vmatrix}$ equal $5 \cdot 6 \cdot 7 \cdot 8$.

10. What can you conclude about $|A|$ if

(a) $A^2 = 0$ (the zero matrix)

(b) $A^2 = I$

11. What can you conclude about $|M|$ if

(a) the reduced echelon form of M is $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

(b) M row operates to $\begin{bmatrix} 2 & 0 & 0 \\ 3 & 4 & 0 \\ 5 & 6 & 7 \end{bmatrix}$

12. Let $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & 1 & 1 & 1 \\ b & a & 1 & 1 \\ c & b & a & 1 \end{bmatrix}$.

Use row or col ops to triangularize A and then find $|A|$

13. Pull out all the factors you can from the rows and cols of $\begin{vmatrix} 1 & 3 & 6 \\ 2 & 4 & 4 \\ 1 & 2 & 8 \end{vmatrix}$.

14. What happens to $|M|$ if row 3 is replaced by 4 row 3 + row 2 (be careful).

15. Find

(a) $\begin{vmatrix} a & b & 0 & 0 \\ 0 & a & b & 0 \\ 0 & 0 & a & b \\ b & 0 & 0 & a \end{vmatrix}$ (b) $\begin{vmatrix} a & b & 0 & 0 & 0 \\ 0 & a & b & 0 & 0 \\ 0 & 0 & a & b & 0 \\ 0 & 0 & 0 & a & b \\ b & 0 & 0 & 0 & a \end{vmatrix}$

16. Let A be a 4×4 matrix.

(a) Suppose B is obtained from A by tripling every entry in A. Write an equation relating
(i) A and B
(ii) $|A|$ and $|B|$

(b) Suppose B is obtained from A by tripling row 2. Write an equation relating
(i) A and B
(ii) $|A|$ and $|B|$

SECTION 1.4 TRANSPOSES

definition of the transpose

To transpose a matrix, turn the rows into cols (and the cols into rows).

$$\text{If } M = \begin{bmatrix} 2 & 3 & 6 \\ 5 & 8 & 9 \end{bmatrix} \text{ then } M^T = \begin{bmatrix} 2 & 5 \\ 3 & 8 \\ 6 & 9 \end{bmatrix}$$

In general the ij-th entry in A^T is the ji-th entry in A.

properties of transposing

(1) $(A + B)^T = A^T + B^T$

(2) $(cA)^T = c\,A^T$

(3) $(AB)^T = B^T A^T$

$(ABC)^T = C^T B^T A^T$ etc.

warning $(ABC)^T$ is *not* $A^T B^T C^T$. Instead the factors get reversed and $(ABC)^T$ is $C^T B^T A^T$.

As a special case, $(A^n)^T = (A^T)^n$.

(4) $(A^T)^T = A$

(5) $|A^T| = |A|$

semiproof of (3)

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } B = \begin{bmatrix} p \\ q \end{bmatrix}$$

Then

$$AB = \begin{bmatrix} ap+bq \\ cp+dq \end{bmatrix}$$

$$(AB)^T = \begin{bmatrix} ap+bq & cp+dq \end{bmatrix}$$

$$B^T A^T = \begin{bmatrix} p & q \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} ap+bq & cp+dq \end{bmatrix}$$

so $(AB)^T = B^T A^T$

dull proof of (5)

For the 2×2 case let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ and $|A|$ and $|A^T|$ are both $ad - bc$. So it is true that $|A| = |A^T|$ in the 2×2 case.

Look at the 3×3 case. Let $A = \begin{bmatrix} a1 & a2 & a3 \\ b1 & b2 & b3 \\ c1 & c2 & c3 \end{bmatrix}$. Then $A^T = \begin{bmatrix} a1 & b1 & c1 \\ a2 & b2 & c2 \\ a3 & b3 & c3 \end{bmatrix}$

and

$$|A| = a1 \begin{vmatrix} b2 & b3 \\ c2 & c3 \end{vmatrix} - a2 \begin{vmatrix} b1 & b3 \\ c1 & c3 \end{vmatrix} + a3 \begin{vmatrix} b1 & b2 \\ c1 & c2 \end{vmatrix} \qquad \text{(expand across row 1)}$$

$$|A^T| = a1 \begin{vmatrix} b2 & c2 \\ b3 & c3 \end{vmatrix} - a2 \begin{vmatrix} b1 & c1 \\ b3 & c3 \end{vmatrix} + a3 \begin{vmatrix} b1 & c1 \\ b2 & c2 \end{vmatrix} \qquad \text{(expand down col 1)}$$

The corresponding 2×2 dets are transposes so they're equal (as just proved).
So $|A| = |A^T|$ in the 3×3 case

Look at the 4×4 case.

$$|A| = \begin{vmatrix} a1 & a2 & a3 & a4 \\ b1 & b2 & b3 & b4 \\ c1 & c2 & c3 & c4 \\ d1 & d2 & d3 & d4 \end{vmatrix} = a1 \begin{vmatrix} b2 & b3 & b4 \\ c2 & c3 & c4 \\ d2 & d3 & d4 \end{vmatrix} - a2 \begin{vmatrix} b1 & b3 & b4 \\ c1 & c3 & c4 \\ d1 & d3 & d4 \end{vmatrix} + a3 \begin{vmatrix} b1 & b2 & b4 \\ c1 & c2 & c4 \\ d1 & d2 & d4 \end{vmatrix} - a4 \begin{vmatrix} b1 & b2 & b3 \\ c1 & c2 & c3 \\ d1 & d2 & d3 \end{vmatrix}$$

$$|A^T| = \begin{vmatrix} a1 & b1 & c1 & d1 \\ a2 & b2 & c2 & d2 \\ a3 & b3 & c3 & d3 \\ a4 & b4 & c4 & d4 \end{vmatrix} = a1 \begin{vmatrix} b2 & c2 & d2 \\ b3 & c3 & d3 \\ b4 & c4 & d4 \end{vmatrix} - a2 \begin{vmatrix} b1 & c1 & d1 \\ b3 & c3 & d3 \\ b4 & c4 & d4 \end{vmatrix} + a3 \begin{vmatrix} b1 & c1 & d1 \\ b2 & c2 & d2 \\ b4 & c4 & d4 \end{vmatrix} - a4 \begin{vmatrix} b1 & c1 & d1 \\ b2 & c2 & d2 \\ b3 & c3 & d3 \end{vmatrix}$$

The corresponding 3×3 dets are transposes so they're equal (as just proved).
So $|A| = |A^T|$ in the 4×4 case.
And so on. Once you've proved it for one size matrix, it's destined to hold for the next size matrix too.

PROBLEMS FOR SECTION 1.4

1. Find the transpose of A^TBC in terms of A,B,C and their transposes.

2. If $B = AA^T$ express $|B|$ in terms of $|A|$.

3. Let A be $r \times n$ and let B be $q \times n$ so that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & \square \\ a_{21} & & & \\ \vdots & & & \\ \square & & & \square \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & \square \\ b_{21} & & & \\ \vdots & & & \\ \square & & & \square \end{bmatrix}$$

(a) Fill in the 3 corner entries in each matrix in correct mathematical notation (like the other entries).

(b) Find the ij-th entry (the entry in the i-th row and j-th col) in BA^T.

SECTION 1.5 INVERSES

definition of the inverse of a square matrix

The inverse of a square matrix is another square matrix called A^{-1} with the property that

$$AA^{-1} = A^{-1}A = I$$

In other words, if $PQ = QP = I$ then P and Q are inverses of each other.

Some square matrices have inverses and some don't. The ones which do are called *invertible* or *nonsingular*.

using the inverse to solve a matrix equation

Look at the equation

$$AX = B$$

Suppose A is invertible. To solve for X, *left* multiply on each side by A^{-1} to get

$$A^{-1}AX = A^{-1}B$$
$$IX = A^{-1}B$$
$$X = A^{-1}B$$

Similarly, to solve

$$XA = B$$

for X, if A is invertible, *right* multiply by A^{-1} to get

$$X = BA^{-1}$$

example 1

Here's how to solve

$$ABCD + E = PQR \quad \text{(all same size square matrices)}$$

for B, assuming that A,C,D are invertible.

$$ABCD = PQR - E$$

$$BCD = A^{-1}(PQR - E) \qquad \text{(left multiply by } A^{-1}\text{)}$$

$$BC = A^{-1}(PQR - E)D^{-1} \qquad \text{(right multiply by } D^{-1}\text{)}$$

$$B = A^{-1}(PQR - E)D^{-1}C^{-1} \qquad \text{(right multiply by } C^{-1}\text{)}$$

computing inverses using the adjoint matrix

(a) If $\det A \neq 0$ then A is invertible and A^{-1} can be painfully computed as follows.

Replace each entry of A by its cofactor and then transpose; the resulting matrix is called the *adjoint* of A. Finally, divide by $|A|$ and you get A^{-1}. In other words,

$$A^{-1} = \frac{\text{adjoint } A}{\det A}$$

Here's an example. Let

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

Then

$$(*)\qquad A^{-1} = \frac{1}{|A|}\underbrace{\begin{bmatrix} +\begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 3 & 2 \end{vmatrix} & +\begin{vmatrix} 0 & 2 \\ 3 & 0 \end{vmatrix} \\ -\begin{vmatrix} 4 & 0 \\ 0 & 2 \end{vmatrix} & +\begin{vmatrix} 2 & 0 \\ 3 & 2 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 3 & 0 \end{vmatrix} \\ +\begin{vmatrix} 4 & 0 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} & +\begin{vmatrix} 2 & 4 \\ 0 & 2 \end{vmatrix} \end{bmatrix}^T}_{\text{adjoint of A}}$$

$$= \frac{1}{20}\begin{bmatrix} 4 & 3 & -6 \\ -8 & 4 & 12 \\ 4 & -2 & 4 \end{bmatrix}^T$$

$$= \begin{bmatrix} 1/5 & -2/5 & 1/5 \\ 3/20 & 1/5 & -1/10 \\ -3/10 & 3/5 & 1/5 \end{bmatrix}$$

warning

When you compute the *determinant* of A by expanding say across row 1 the first term in the sum is the entry 2 *times* its cofactor:

$$2\cdot\begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix}$$

But when you compute the *adjoint* to get the inverse there is no "2 times". The entries in the adjoint are plain cofactors, not cofactors *times* anything.

(b) If det A = 0 then A is not invertible

(Naturally the particular formula in (a) can't hold because you can't divide by 0 but furthermore no formula of any kind can hold for A^{-1} because there is no A^{-1}.)

Put (a) and (b) together:

$$\boxed{\text{A is invertible if and only if } |A| \neq 0}$$

how to write a "proof by contradiction"

If you want to prove by contradiction that X is red then the argument must have this form:

Suppose X is not red.
⋮
xxxx
But xxxx is impossible
So X must be red.

proof of (b)

Let $|A| = 0$. Suppose A does have an inverse. I'll get a contradiction.

$$AA^{-1} = I$$
$$|AA^{-1}| = |I|$$
$$|A||A^{-1}| = 1$$

But $|A| = 0$ so the product $|A||A^{-1}|$ can't be 1. Contradiction.
So A can't have an inverse.

proof of (a) (pretty messy)

I'll do it for a 5×5 matrix. Let

$$A = \begin{bmatrix} a1 & a2 & a3 & a4 & a5 \\ b1 & b2 & b3 & b4 & b5 \\ c1 & c2 & c3 & c4 & c5 \\ d1 & d2 & d3 & d4 & d5 \\ e1 & e2 & e3 & e4 & e5 \end{bmatrix}$$

Remember that det A is found by going across any row (or col), multiplying each entry by its cofactor and adding. For example

$$\det A = e1 \cdot \text{cof}\, e1 + e2 \cdot \text{cof}\, e2 + e3 \cdot \text{cof}\, e3 + e4 \cdot \text{cof}\, e4 + e5 \cdot \text{cof}\, e5$$

To half-prove that the formula in (1) does give the inverse, I'll find

$$(1) \qquad A \times \underbrace{\frac{1}{|A|} \begin{bmatrix} \text{cof}\, a1 & \text{cof}\, a2 & \text{cof}\, a3 & \text{cof}\, a4 & \text{cof}\, a5 \\ \text{cof}\, b1 & \text{cof}\, b2 & \text{cof}\, b3 & \text{cof}\, b4 & \text{cof}\, b5 \\ \text{cof}\, c1 & \text{cof}\, c2 & \text{cof}\, c3 & \text{cof}\, c4 & \text{cof}c5 \\ \text{cof}\, d1 & \text{cof}\, d2 & \text{cof}\, d3 & \text{cof}\, d4 & \text{cof}\, d5 \\ \text{cof}\, e1 & \text{cof}\, e2 & \text{cof}\, e3 & \text{cof}\, e4 & \text{cof}\, e5 \end{bmatrix}^T}_{\text{supposed inverse of A}}$$

and show that it is I. Multiply out in (1) (don't forget to transpose the matrix of cofactors):

1,1 entry of the product in (1)

$$= \frac{1}{|A|} \underbrace{(a1 \cdot \text{cof}\, a1 + a2 \cdot \text{cof}\, a2 + a3 \cdot \text{cof}\, a3 + a4 \cdot \text{cof}\, a4 + a5 \cdot \text{cof}\, a5)}_{\text{This is det A computed by expanding across row 1 of A}}$$

$$= \frac{1}{|A|} \times |A|$$
$$= 1$$

Similarly, the 2,2 entry, the 3,3 entry, the 4,4 entry and the 5,5 entry in the product in (1) are all 1.

Now look at the entries in the product in (1) that are *off* the main diagonal.

4,1 entry of the product in (1)

$$(2) \qquad = \frac{1}{|A|} \underbrace{(d1 \cdot \text{cof}\, a1 + d2 \cdot \text{cof}\, a2 + d3 \cdot \text{cof}\, a3 + d4 \cdot \text{cof}\, a4 + d5 \cdot \text{cof}\, a5)}_{\text{Not det A}}$$

The second factor in (2) is *not* det A because the d's are not multiplied by their *own* cofactors. But it happens to be the determinant of the matrix

$$(3) \qquad \begin{bmatrix} d1 & d2 & d3 & d4 & d5 \\ b1 & b2 & b3 & b4 & b5 \\ c1 & c2 & c3 & c4 & c5 \\ d1 & d2 & d3 & d4 & d5 \\ e1 & e2 & e3 & e4 & e5 \end{bmatrix}$$

And the determinant of (3) is 0 because it has two identical rows. So all in all,

4,1 entry of the product in (1) = det of (3) = 0

Similarly all the entries off the main diagonal of the product in (1) are 0. So the product in (1) is I.

For the other half of the proof, a similar argument can be given to show that when you multiply in the opposite order,

$$(4) \qquad \underbrace{\frac{1}{|A|}\begin{bmatrix} \text{cof } a1 & \text{cof } a2 & \text{cof } a3 & \text{cof } a4 & \text{cof } a5 \\ \text{cof } b1 & \text{cof } b2 & \text{cof } b3 & \text{cof } b4 & \text{cof } b5 \\ \text{cof } c1 & \text{cof } c2 & \text{cof } c3 & \text{cof } c4 & \text{cof } c5 \\ \text{cof } d1 & \text{cof } d2 & \text{cof } d3 & \text{cof } d4 & \text{cof } d5 \\ \text{cof } e1 & \text{cof } e2 & \text{cof } e3 & \text{cof } e4 & \text{cof } e5 \end{bmatrix}^T}_{\text{supposed inverse of A}} \times A,$$

you also get I. Put the two halves together and you know that the supposed inverse of A *is* the inverse of A. QED

The key idea turned out to be this obscure fact: If you "expand" across a row of a matrix A and use the cofactors of that row the result is det A but if you "expand" across one row and use the cofactors that belong to *another* row the result is 0.

the adjoint method in the special case of 2 × 2 matrices

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

If $|A| = 0$ then A has no inverse. If $|A| \neq 0$ then

$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} +d & -c \\ -b & +a \end{bmatrix}^T \quad \text{by} \quad (1)$$

$$= \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

(5) $\boxed{\text{To find the inverse of a } 2 \times 2 \text{ matrix (if the det is nonzero), interchange the entries on the main diagonal, change the signs of the other two entries, and divide by the determinant.}}$

If $A = \begin{bmatrix} 2 & 3 \\ 5 & 1 \end{bmatrix}$ then $A^{-1} = \frac{1}{-13}\begin{bmatrix} 1 & -3 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} -1/13 & 3/13 \\ 5/13 & -2/13 \end{bmatrix}$

If $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ then A doesn't have an inverse since $|A| = 0$

warning

The method in (5) is the best one for inverting 2 × 2's. *Use it.*

But don't forget to divide by the determinant.

the one-sided rule for testing for inverses

To see if A and B are inverses it's enough to see if AB = I. If it is then A and B are inverses and it isn't necessary to check further to see if BA = I because it always will be.

proof

Suppose AB = I. I'm going to show that BA = I also.

First I'll show that A has an inverse.
If AB = I then $|AB| = |I|$, $|A||B| = 1$ so $|A| \neq 0$. So A is invertible.
Now that I know A^{-1} exists I can use it to turn BA into I.

$$
\begin{aligned}
BA &= A^{-1}A\,BA && (A^{-1}A \text{ is I and multiplying by I doesn't change anything})\\
&= A^{-1}\,I\,A && (\text{the AB in the middle is I, by hypothesis})\\
&= A^{-1}A\\
&= I && \text{QED}
\end{aligned}
$$

the inverse of a diagonal matrix

Let $A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$.

(a) If a,b,c are all nonzero then $A^{-1} = \begin{bmatrix} 1/a & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1/c \end{bmatrix}$

(b) If any one of a,b,c is zero then A is noninvertible.

proof of (a)

Just multiply the two matrices to see that you get I.

proof of (b)

Remember that the det of a diagonal matrix is the product of the diagonal entries. So if one of a,b,c is 0 then $|A| = 0$ so A has no inverse.

computing inverses using row ops (another method for inverting)

To get the inverse of A, try to row operate A until it turns into I.

(a) If A can't row operate into I then A has no inverse.

(b) If A can be row opped to I then A has an inverse and performing the same row ops on I will turn I into A^{-1}.

Here's an example. Let

$$A = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 0 & 0 \\ 0 & -2 & 2 \end{bmatrix}$$

To find A^{-1}, if it exists, start with A and I and row operate as follows:

$$\begin{array}{lccc}
 & \begin{matrix} 0 & 1 & 4 \\ 1 & 0 & 0 \\ 0 & -2 & 2 \end{matrix} & \qquad & \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \\[2ex]
R_1 \leftrightarrow R_2 & \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & -2 & 2 \end{matrix} & & \begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{matrix} \\[2ex]
R_3 = 2R_2 + R_3 & \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 10 \end{matrix} & & \begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 2 & 0 & 1 \end{matrix} \\[2ex]
R_3 = \frac{1}{10} R_3 & \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{matrix} & & \begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1/5 & 0 & 1/10 \end{matrix} \\[2ex]
R_2 = -4R_3 + R_2 & \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} & & \underbrace{\begin{matrix} 0 & 1 & 0 \\ 1/5 & 0 & -2/5 \\ 1/5 & 0 & 1/10 \end{matrix}}_{\text{This is } A^{-1}.}
\end{array}$$

proof of (a)

If A won't row operate into I then the echelon form must have a col without a pivot. For example, the echelon form looks something like

$$\begin{matrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{matrix} \quad \text{or} \quad \begin{matrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix} \quad \text{or} \quad \begin{matrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix} \quad \text{etc.}$$

But then the echelon form will have an all zero row and its determinant is 0. Row ops don't change the zero-ness of a det so the original det is 0 also. So A doesn't have an inverse.

proof of (b)

Suppose it takes 3 row ops to turn A into I. Each row op has a corresponding elementary matrix; call them E_1, E_2, E_3. Then the procedure can be described like this:

start with A and I	A	I
left mult by E_1 (i.e., do the first row op)	E_1A	E_1I
left mult by E_2 (i.e., do the second row op)	E_2E_1A	E_2E_1I

left mult by E_3 (i.e., do the third row op) $\underbrace{E_3E_2E_1A}_{\text{This is I.}}$ $\underbrace{E_3E_2E_1I}_{\text{This is } E_3E_2E_1.}$

The row ops turn A into I so $E_3E_2E_1A = I$.

But that makes $E_3E_2E_1$ and A inverses (by the one-sided above). So $A^{-1} = E_3E_2E_1$.
But what you have in the righthand column is precisely $E_3E_2E_1$. So doing the row ops to I does turn I into A^{-1}.

example 2

I'll use row ops to find the inverse of $M = \begin{bmatrix} 2 & -2 & 8 \\ 3 & 1 & 12 \\ 9 & 1 & 36 \end{bmatrix}$.

Start with M and I.

	$\begin{matrix} 2 & -2 & 8 \\ 3 & 1 & 12 \\ 9 & 1 & 36 \end{matrix}$	$\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$
$R_1 = \frac{1}{2} R_1$	$\begin{matrix} 1 & -1 & 4 \\ 3 & 1 & 12 \\ 9 & 1 & 36 \end{matrix}$	$\begin{matrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$
$R_2 = -3R_1 + R_2$ $R_3 = -9R_1 + R_3$	$\begin{matrix} 1 & -1 & 4 \\ 0 & 4 & 0 \\ 0 & 10 & 0 \end{matrix}$	$\begin{matrix} 1/2 & 0 & 0 \\ -3/2 & 1 & 0 \\ -9/2 & 0 & 1 \end{matrix}$
$R_2 = \frac{1}{4} R_2$	$\begin{matrix} 1 & -1 & 4 \\ 0 & 1 & 0 \\ 0 & 10 & 0 \end{matrix}$	$\begin{matrix} 1/2 & 0 & 0 \\ -3/8 & 1/4 & 0 \\ -9/2 & 0 & 1 \end{matrix}$
$R_2 = R_1 + R_2$ $R_3 = -10R_2 + R_3$	$\begin{matrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{matrix}$	Forget it.

It's impossible to continue on the left side to get I .
So M has no inverse and all the work done on the right side was wasted.

invertible rule

Let A be a square matrix.
The following are equivalent, i.e., either all are true or all are false.

(1) A is invertible (i.e., non-singular).
(2) $|A| \neq 0$.
(3) Reduced echelon form of A is I.

The equivalence of (1) and (2) comes from the adjoint method for inverting.
The equivalence of (1) and (3) comes from the row op method for inverting.

The invertible rule says a whole lot. For instance it says:
If $|A| \neq 0$ then A is invertible.
If A is invertible then $|A| \neq 0$.
If $|A| = 0$ then A is not invertible.
If A is not invertible then $|A| = 0$.
If A can be row operated to I then A is invertible.
If $|A| = 0$ then A can't be row operated to I.
And so on.

algebra of inverses

(1) If A is invertible then so is A^{-1} and in that case

$$\boxed{(A^{-1})^{-1} = A}$$

In other words, A and A^{-1} are inverses of each other.

(2) If A is invertible then so is A^T and

$$\boxed{(A^T)^{-1} = (A^{-1})^T}$$

(3) If A,B,C are invertible then so is ABC and in that case

$$\boxed{(ABC)^{-1} = C^{-1}B^{-1}A^{-1}}$$

warning $(ABC)^{-1}$ is *not* $A^{-1}B^{-1}C^{-1}$. Instead the factors get reversed and you get $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.

In particular, if A is invertible then so is A^n and in that case

$$\boxed{(A^n)^{-1} = (A^{-1})^n}$$

(4) If A is invertible and c is a nonzero number then cA is invertible and in that case

$$\boxed{(cA)^{-1} = \frac{1}{c} A^{-1}}$$

For example, if A is invertible then 3A is also invertible and $(3A)^{-1} = \frac{1}{3} A^{-1}$.

proof of (4)

If you think you have an inverse for a matrix Q, check it out by multiplying the two together to see if you get I.

In particular, given that A is invertible, to show that the inverse of cA is $\frac{1}{c} A^{-1}$, multiply them together to see if you get I.

$$\begin{aligned} \frac{1}{c} A^{-1} cA &= \frac{1}{c} c A^{-1}A \quad \text{(by property (7) of scalar mult; Section 1.1)} \\ &= A^{-1}A \\ &= I \quad \text{QED} \end{aligned}$$

This not only shows that if A is invertible then cA is invertible, it also shows that the inverse is $\frac{1}{c} A^{-1}$ (by the one-sided rule for testing for inverses).

Here's another way to show that if A is invertible then cA is invertible, but it's only half a proof of (4) because it doesn't produce the actual inverse of cA. By a rule of determinants,

$$|cA| = c^n|A|$$

But $|A| \neq 0$ since A is invertible and $c^n \neq 0$ since $c \neq 0$ so $|cA| \neq 0$. So cA is invertible.

proof of (2)

method 1 (only half a proof) I'll show that if A is invertible then so is A^T.

If A is invertible then $|A| \ne 0$. But $|A^T| = |A|$ so $|A^T| \ne 0$. But that makes A^T invertible. But this method hasn't actually found what the inverse is.

method 2 I'll show that if A is invertible then $(A^{-1})^T$ and A^T are inverses. To check that they are inverses, multiply them together:

$$\begin{aligned} A^T(A^{-1})^T &= (A^{-1}A)^T \text{ by the transpose rule } (PQ)^T = Q^TP^T \\ &= I^T \\ &= I \quad \text{QED} \end{aligned}$$

This not only shows that A^T is invertible, it also shows that the inverse is $(A^{-1})^T$.

proof of (3)

I'll show that $(ABC)(C^{-1}B^{-1}A^{-1}) = I$. That not only shows that ABC is invertible, it shows that the inverse is $C^{-1}B^{-1}A^{-1}$.

$$\begin{aligned} (ABC)(C^{-1}B^{-1}A^{-1}) &= A\,B\,\underbrace{C\,C^{-1}}_{I}\,B^{-1}\,A^{-1} \qquad \text{(I can group or ungroup as I choose)} \\ &= A\,\underbrace{B\,B^{-1}}_{I}\,A^{-1} \\ &= AA^{-1} \\ &= I \qquad \text{QED} \end{aligned}$$

determinant of the inverse

If A is invertible then

$$\boxed{|A^{-1}| = \frac{1}{|A|}}$$

For example, if $|A| = \frac{3}{4}$ then $|A^{-1}| = \frac{4}{3}$.

proof

$$AA^{-1} = I$$

$$|AA^{-1}| = |I|$$

$$|A||A^{-1}| = 1$$

$$|A^{-1}| = \frac{1}{|A|}$$

warning

1. Non-square matrices don't have inverses. Even some square matrices don't have inverses. You can't take inverses for granted.

2. $(A + B)^{-1}$ is not $A^{-1} + B^{-1}$. It is true that $(A + B)^T$ is $A^T + B^T$ but it doesn't work like that for inverses.

3. The algebra rule (3) does *not* say $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.
It says that *if* A,B,C are invertible *then* $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.
You can't inverse out a product unless you know in advance that the individual factors are invertible.

PROBLEMS FOR SECTION 1.5

1. It is not necessarily true that if AB = AC then A cancels to leave B = C. But show that it *is* true if A is invertible.

2. Suppose PQR = C (all matrices same size) where P,R are invertible. Solve for Q.

3. Let EABF + CD = I (all matrices same size square). Decide what inverses you need to be able to solve for B and then, assuming those inverses exist, do the solving.

4. Find (a) $ABC(ABC)^{-1}$ (b) $(2A - B)(2A - B)^{-1}$ (c) $(3A)(3A)^{-1}$

5. Suppose A and B commute, i.e., AB = BA, and A is invertible.
Show that $(A^{-1}B^2A)^2 = B^4$.

6. Use the adjoint method to find inverses if they exist.

(a) $\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

7. Let A be the matrix in #6(c).

Find a column vector $\vec{x}$ so that $A\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

8. Let

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 6 & 8 & 10 & 12 \end{bmatrix}$$

Find the entry in row 3, col 2 of A^{-1}.

9. Let A = [4], a 1×1 matrix. Find its inverse.

10. Find the inverse of $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

11. Show that a (square) matrix can't have two inverses (either it has none or it has one). Here's the beginning and end of the proof. You fill in the middle.

Suppose AB = BA = I so that B is an inverse for A.
And suppose that AC = CA = I so that C is also an inverse for A.

⋮

Then B = C so the two inverses are one and the same.

12. For what value(s) of x is $\begin{bmatrix} x-3 & 4 \\ 2 & x-1 \end{bmatrix}$ invertible.

13. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Find B if $BA = \begin{bmatrix} 1 & 2 \\ 4 & 6 \\ 0 & 2 \end{bmatrix}$.

14. Find the inverse using row ops.

(a) $A = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

15. What happens if you try to invert $\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$ using row ops.

16. Simplify (a) $(A^{-1})^{-1}$ (b) $(A^T)^T$ (c) $(A^3)^3$

17. Get rid of the parentheses.
(a) $(AB)^T$ (b) $(AB)^{-1}$ assuming that A and B are invertible (c) $(AB)^3$

18. Assume all inverses exist. Simplify (get rid of all parentheses and brackets).
(a) $(A^{-1}B)^{-1}$ (b) $[\, A(A+B)^{-1}\,]^{-1}$

19. Assume all inverses exist. Suppose you know the inverse of A^TA ; call it Q. Express the inverse of A itself in terms of Q.

20. Suppose $A^{-1} = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 0 & 0 & 1 \end{bmatrix}$.

Find (a) $(A^T)^{-1}$ (b) $(2A)^{-1}$ (c) $|A|$ (d) $|2A|$

21. If $|A| = -3$ find (a) $|4A^{-1}|$ (b) $|(4A)^{-1}|$

22. Assume A and B are the same size square.
We have this rule: If A and B are invertible then AB is invertible.
Show that the converse is true: If AB is invertible then A and B are invertible.

23. Assume A and B are the same size square.
We have this rule: If A and B are invertible then AB is invertible.
Suppose A is invertible but B isn't. What about AB?

24. Suppose A is invertible. You invert A. Your friend inverts A^T. How are your answers related.

25. Find $|A^{-1}|$ if $A = \begin{bmatrix} 1/7 & 2 & 6 \\ \pi & 1 & 2 \\ 4 & 0 & 0 \end{bmatrix}$.

26. True or False: If A and B are invertible then A+B is also invertible.
If true, prove it. If false, find a counterexample.

SECTION 1.6 SYMMETRIC AND SKEW SYMMETRIC MATRICES

symmetric matrices

A square matrix M (with real entries) is called symmetric if $M^T = M$. In other words, a matrix is symmetric if the "matching" entries a_{ij} and a_{ji} are equal.

For example, $\begin{bmatrix} 1 & 5 & 6 \\ 5 & 0 & -2 \\ 6 & -2 & 3 \end{bmatrix}$ is symmetric.

> In general, to show that a matrix is symmetric:
>
> *method 1* Look at the matrix to see if its matching entries are equal.
>
> *method 2* To show that say $AB + CD^{-1}$ is symmetric, given some hypotheses, find $(AB + CD^{-1})^T$ using matrix algebra and show that it simplifies to $AB + CD^{-1}$.
>
> Your argument should look something like this:
>
> $$\begin{aligned} (AB + CD^{-1})^T &= \ldots \\ &= \ldots \\ &= \ldots \\ &= AB + CD^{-1} \end{aligned}$$
>
> Therefore $AB + CD^{-1}$ is symmetric.

example 1

Show that if A is symmetric and invertible then A^{-1} is also symmetric.

method 1 Start with an abstract $n \times n$ matrix that is symmetric and invertible. Compute it's inverse and look at it to see if its matching entries are equal. This can be messy (see problem 7).

method 2 (better) Let A be symmetric and invertible. I'll show that

$$(A^{-1})^T = A^{-1}. \tag{$*$}$$

To do this, I'll start with the left side of (*) and work on it until it turns into A^{-1}:

$$\begin{aligned} (A^{-1})^T &= (A^T)^{-1} && \text{(inverse rule)} \\ &= A^{-1} && \text{(since A is symmetric)} \quad \text{QED} \end{aligned}$$

Therefore A^{-1} is symmetric.

warning about style

To show that $(A^{-1})^T = A^{-1}$ in example 1, it is neither good style nor good logic to write like this:

Don't write like this	$(A^{-1})^T = A^{-1}$	(what you want to prove)	*Don't write like this*
	$(A^T)^{-1} = A^{-1}$	(inverse rule)	
	$A^{-1} = A^{-1}$	(since A is symmetric)	
	TRUE!		

Any "proof" in mathematics that *begins* with what you want to prove and *ends* with something TRUE, like $A^{-1} = A^{-1}$ (or $B = B$ or $0 = 0$) is at best badly written and at worst incorrect and *drives me crazy*.

What you *should* do to prove that $(A^{-1})^T$ equals A^{-1} is work on one of them until it turns into the other or work on each one *separately* until they turn into the same thing. Don't write $(A^{-1})^T = A^{-1}$ as the *first* line of your proof. It should be your *last* line.

skew symmetric matrices

A square matrix (with real entries) is *skew symmetric* if $M^T = -M$ or equivalently if $-M^T = M$. In other words, a matrix is skew-symmetric if the diagonal entries are 0 and the matching entries a_{ij} and a_{ji} are negatives of one another.

For example, $\begin{bmatrix} 0 & 3 & 4 \\ -3 & 0 & -5 \\ -4 & 5 & 0 \end{bmatrix}$ is skew symm.

warning
Don't forget that a skew symm must have 0's on the main diagonal.

In general, to show that a matrix is skew symm:

method 1 Look at M to see if its diagonal entries are 0 and its matching entries negatives of one another.

method 2 To show that say $AB + CD^{-1}$ is symmetric, given some hypotheses, find $(AB + CD^{-1})^T$ using matrix algebra and show that it simplifies to $AB + CD^{-1}$.

PROBLEMS FOR SECTION 1.6

1. Show that $A + A^T$ is symmetric for every square A. Do it twice.
(a) Use (1) above for the 3×3 case (write out an arbitrary 3×3 matrix A, find $A + A^T$ and look at it).
(b) Do it for the $n \times n$ case in general, using (2) above.

2. (a) Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. Compute AA^T to see that it's symmetric

(b) Show that AA^T is symmetric for any $n \times m$ matrix in general.

3. Show that if A is symmetric then B^TAB is also symmetric.

4. True or False. If false, produce a counterexample.

(a) The sum of symmetric matrices is symm.
(b) If A is symm then 5A is also symm.
(c) The product of symm matrices is symm.

5. Example 1 showed that if A is invertible and symmetric then A^{-1} is symmetric. Do it again, say in the 4×4 case, by actually computing A^{-1} (as much as necessary) by the adjoint method.

6. Show that if A is skew symm then A and A^T commute (meaning that AA^T is the same as A^TA).

7. If A is skew symmetric is it true that A^6 must also be skew symm.

REVIEW PROBLEMS FOR CHAPTER 1

1. Find D^9 if If D is a 10×10 diagonal matrix with diagonal entries $d_1, \ldots, d_{10}$.

2. Find the inverse of $\begin{bmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 4 & 0 \\ 0 & 5 & 0 & 0 \\ 6 & 0 & 0 & 0 \end{bmatrix}$.

3. Look at the following statement: If A is an $n \times n$ skew symm matrix then $|A| = 0$.
 (a) Find a counterexample to disprove the statement in the case that n is even; i.e., produce one specific skew symm matrix, either 2×2 or 4×4 or 6×6 etc. whose det is not 0.
 (b) Show that the statement is true if $n = 3$ by writing out an arbitrary 3×3 skew symm matrix and computing its determinant.
 (c) Go further than part (b)and show that the statement is true for *any* odd n.
 (It's too messy to write out an arbitrary $n \times n$ skew symm matrix matrix and compute its det. Be more subtle and use general det theory.)

4. Show that if $A^2 = 0$ then A is not invertible.

5. The trace of a square matrix is the sum of its diagonal entries.
 Is it true that trace(AB) = (trace A)(trace B) for all square A,B.
 Defend your answer.

6. (a) Show that it is possible for both A and B to be non-zero and still have a zero product.
 (b) But show that it is not possible in the special case that A and B are invertible.

7. If A is skew symm is there anything special about $A + A^T$.

8. Suppose $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 8$. Find

(a) $\begin{vmatrix} a & -b & c \\ d & -e & f \\ g & -h & i \end{vmatrix}$ (b) $\begin{vmatrix} a & b & c \\ d+2a & e+2b & f+2c \\ 3g & 3h & 3i \end{vmatrix}$ (c) $\begin{vmatrix} a & b & c \\ 4d+2a & 4e+2b & 4f+2c \\ 3g & 3h & 3i \end{vmatrix}$

9. Let $A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 3 & 0 & 1 & 2 \\ 2 & 1 & -1 & 1 \end{bmatrix}$. Find the echelon form of A.

10. A and B are both $n \times n$ matrices.
 You know B and you know $(AB)^{-1}$ but you don't know A.
(a) How do you know that A inverts.
(b) How would you find the inverse of A in terms of the matrices you already know.

11. Are the following reduced echelon form, unreduced echelon form or not echelon form at all.

(a) $\begin{matrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{matrix}$ (b) $\begin{matrix} 1 & 0 & 2 \\ 0 & 0 & 3 \\ 0 & 1 & 4 \end{matrix}$ (c) $\begin{matrix} 3 & 0 & 5 \\ 0 & 1 & 6 \\ 0 & 0 & 0 \end{matrix}$ (d) $\begin{matrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix}$

(e) $\begin{matrix} 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 \end{matrix}$ (f) $\begin{matrix} 4 & 5 & 6 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{matrix}$

12. Remember (Section 1.5) that the adjoint of a square matrix is the transpose of the matrix of cofactors.

(a) Show (easily) that if M is an invertible $n \times n$ matrix then adj M is also invertible.
(b) (i) Suppose $|A| = 3$ and A is 7×7. Find $|\text{adj } A|$.
(ii) Now generalize (i). Suppose $|A|$ is nonzero and A is $n \times n$. Express $|\text{adj } A|$ in terms of $|A|$.

13. Find $\begin{vmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{vmatrix}$

(a) by expanding across a row or down a col.
(b) using the definition of the determinant as a sum of signed products

14. Let A be a square matrix.
(a) You find $|A|$. Your lazy roommate is supposed to find $|A^3|$. She takes your answer and cubes it. Is that OK.
(b) You find $|A|$. Your roommate is supposed to find $|3A|$. She takes your answer and triples it. Is she right.
(c) You find the inverse of A. Your roommate is supposed to find the inverse of A^T. She steals your answer and transposes it. Does she get the right answer.
(d) You find the inverse of A. Your roommate is supposed to find the inverse of 2A. She does it by taking your answer and doubling it. Is that right.

15. Let $A = \begin{bmatrix} x+1 & x & 2 & 5 \\ 1 & 1 & x+1 & 2 \\ 0 & 0 & 5 & 2x+1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(a) For what value(s) of x is A invertible.
(b) For those values of x, find $|3A^{-1}|$.

16. Use 3×3 matrices in this problem.
(a) Consider the row operation $R_1 = 2R_3 + R_1$.
Find the "inverse" row op, i.e., the row op that undoes it so that when the two are performed successively the result is the original matrix.

(b) Find the elementary matrices E_1 and E_2 corresponding respectively to the row op $R_1 = 2R_3 + R_1$ and to the opposite row op from (a).
Then check that E_1 and E_2 are (fittingly) inverse matrices.

17. Does $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$ have an inverse.

18. Suppose the matrix A is 10×8 and A^TA is invertible.
Let $Q = A(A^TA)^{-1} A^T$

(a) Why shouldn't you write $(A^T A)^{-1} = A^{-1} (A^T)^{-1}$.
(b) Find Q^2 and get a pretty answer.
(c) Find Q^{100}.
(d) Show that Q is symmetric.

CHAPTER 2 THE VECTOR SPACE R^n

SECTION 2.1 VECTOR OPERATIONS

n-dimensional space

An n-tuple $(u_1,\ldots, u_n)$ is called a vector or point and might be denoted by $\vec{u}$.
(I'll leave out the overhead arrow when I get tired of putting it in.)
The numbers $u_1, \ldots, u_n$ are called the coordinates or the components of the vector.

The vector with all zero components is called the zero vector and denoted by $\vec{0}$.
In the context of a vector discussion, plain numbers are called scalars.
The set of all n-tuples of real numbers is called R^n.

addition, subtraction and scalar multiplication

Addition, subtraction and scalar multiplication is done componentwise:

If $u = (u_1,\ldots, u_n)$, $v = (v_1,\ldots, v_n)$ and c is a scalar (i.e., a number) then

$$u + v = (u_1 + v_1, \ldots, u_n + v_n)$$
$$u - v = (u_1 - v_1, \ldots, u_n - v_n)$$
$$cu = (cu_1,\ldots, cu_n)$$

For example, if

$$u = (1,2,6,3) \text{ and } v = (2,6,7,4)$$

then u and v are in R^4 and

$$u + v = (3,8,13,7)$$

$$u - v = (-1,-4,-1,-1)$$

$$7v = (14,42,49,28)$$

properties of addition and scalar mult

Let a,b be scalars; let u,v,w be vectors.

(1) $u + v = v + u$
(2) $(u + v) + w = u + (v + w)$
(3) $u + \vec{0} = \vec{0} + u = u$
(4) $a(u + v) = au + av$
(5) $au + bu = (a + b)u$; e.g., $2u + 5u = 7u$
(6) $a(bu) = (ab)u$; e.g., $2(3u) = 6u$

the dot product (inner product)

If $u = (u_1,\ldots, u_n)$ and $v = (v_1,\ldots, v_n)$ then

$$u \cdot v = u_1v_1 + \ldots + u_nv_n$$

The dot product $u \cdot v$ is also denoted by $\langle u,v \rangle$.

For example, if $u = (5,4,6,2)$ and $v = (3,2,1,-1)$ then

$$u \cdot v = 15 + 8 + 6 - 2 = 27$$

warning

The dot product is a *scalar*.
In this last example, $u \cdot v$ is 15 *plus* 8 *plus* 6 *plus* -2, not $(15,8,6,-2)$.

properties of the dot product

Let u,v,w,p be vectors in R^n and let k be a scalar.

(1) $u \cdot v = v \cdot u$

(2) $(ku)\cdot v = k(u \cdot v)$

(3) $u \cdot (kv) = k(u \cdot v)$

(4) $u \cdot (v + w) = u \cdot v + u \cdot w$

$(u + v)\cdot(w + p) = u \cdot w + v \cdot p + v \cdot w + u \cdot p$

(5) $u \cdot u > 0$ if $\vec{u} \neq \vec{0}$

$u \cdot u = 0$ if $u = \vec{0}$

In other words, $u \cdot u \geq 0$ and $u \cdot u$ equals 0 only if $u = \vec{0}$.

(6) $u \cdot \vec{0} = 0$

proof of (2)

Let $u = (u_1,\ldots, u_n)$ and $v = (v_1,\ldots, v_n)$. Then

$(ku)\cdot v = (ku_1,\ldots, ku_n)\cdot(v_1,\ldots, v_n)$ by definition of scalar mult

$= ku_1v_1 + \ldots + ku_nv_n$ by definition of the dot product

On the other hand,

$k(u \cdot v) = k(u_1v_1 + \ldots + u_nv_n)$ by definition of the dot product

$= ku_1v_1 + \ldots + ku_nv_n$ by algebra rules for numbers

So $(ku)\cdot v = k(u \cdot v)$

example 1

If $u \cdot v = 3$ and $u \cdot u = 4$ then

$$u \cdot (u + 5v) = u \cdot u + u \cdot 5v = u \cdot u + 5(u \cdot v) = 4 + 5(3) = 19$$

orthogonal (perpendicular) vectors

If $u \cdot v = 0$ then u and v are called *orthogonal*.

For example, if $u = (2,3,4,5)$ and $v = (-\frac{1}{2}, 1, 2, -2)$ then $u \cdot v = 0$ so u and v are orthogonal.

The zero vector in R^n is orthogonal to every vector in R^n.
The vectors u, v, w, p are called orthogonal if each of the vectors is orthog to the other three.

the norm (magnitude) of a vector

$$\text{If } u = (u_1,\ldots, u_n) \text{ then } \|u\| = \sqrt{u_1^2 + \ldots + u_n^2}\,.$$

For example, if $u = (2,3,1,-4)$ then $\|u\| = \sqrt{4 + 9 + 1 + 16} = \sqrt{30}$

properties of norms

Let u be a vector and let k be a scalar.

(1) $u \cdot u = \|u\|^2$

(2) $\|u\| > 0$ if $u \neq \vec{0}$

$\|u\| = 0$ if $u = \vec{0}$

In other words, $\|u\| \geq 0$ and $\|u\| = 0$ only if $u = \vec{0}$.

(3) $\|ku\| = |k|\,\|u\|$ where $|k|$ is the absolute value of the number k.

For example if $\|u\| = 5$ then $-3u$ has norm 15.

(4) (triangle inequality) $\|u + v\| \le \|u\| + \|v\|$ (Fig 4)

proof of (3)

Let $u = (u_1,\ldots,u_n)$. Then $ku = (ku_1,\ldots,ku_n)$ and

$$\|ku\| = \sqrt{(ku_1)^2 + \ldots + (ku_n)^2} \quad \text{definition of the norm}$$

$$= \sqrt{k^2 u_1^2 + \ldots + k^2 u_n^2} \quad \text{algebra}$$

$$= \sqrt{k^2}\sqrt{u_1^2 + \ldots + u_n^2} \quad \text{algebra}$$

$$= |k| \sqrt{u_1^2 + \ldots + u_n^2} \quad \text{algebra}$$

$$= |k|\,\|u\| \quad \text{definition of the norm}$$

example 2

Find $\|u + v\|$ if $\|u\| = 6$, $\|v\| = 7$ and $u\cdot v = 8$.

solution $\|u+v\| = \sqrt{(u+v)\cdot(u+v)}$ by property (1) of norms

$= \sqrt{u\cdot u + 2(u\cdot v) + v\cdot v}$ by property (4) of dots

$= \sqrt{\|u\|^2 + 2(u\cdot v) + \|v\|^2}$ by property (1) of norms

$= \sqrt{36 + 16 + 49}$

$= \sqrt{101}$

warning

1. There is no such thing as u^2. There's a dot product, $u\cdot u$, and a norm, $\|u\|$, and a norm squared, $\|u\|^2$, but nothing is denoted u^2.
2. The dot product is written as $u\cdot v$ or <u,v> but *not* uv. Don't leave out the dot.

normalizing a vector

If $\|v\| = 1$ then v is called a *unit vector* or a *normalized vector*.
Note that

$$\boxed{\text{if } v \text{ is a unit vector then } v\cdot v = 1}$$

since $v\cdot v = \|v\|^2$.

The vector $\dfrac{u}{\|u\|}$ is a unit vector. I'll call it u_{unit}; it's also called the normalized u. In other words, if $u = (u_1,\ldots,u_n)$ then

$$u_{\text{unit}} = \frac{u}{\|u\|} = \left(\frac{u_1}{\|u\|}, \ldots, \frac{u_n}{\|u\|}\right)$$

For example, if $u = (1,4,2,3)$ then $u_{\text{unit}} = \left(\frac{1}{\sqrt{30}}, \frac{4}{\sqrt{30}}, \frac{2}{\sqrt{30}}, \frac{3}{\sqrt{30}}\right)$.

The process of finding u_{unit} is called *normalizing* u.

orthonormal vectors

If $\vec{u}_1, \dots, \vec{u}_n$ are orthogonal unit vectors, they are called *orthonormal*.

n-dimensional distance

The distance between $\vec{x}$ and $\vec{y}$ is defined to be $\|\vec{x}-\vec{y}\|$ or equivalently $\|\vec{y}-\vec{x}\|$.

So if $\vec{x} = (x_1, \dots, x_n)$ and $\vec{y} = (y_1, \dots, y_n)$ then

$$\text{distance between } \vec{x} \text{ and } \vec{y} = \sqrt{(x_1-y_1)^2 + \dots + (x_n-y_n)^2}$$

For example, if $x = (1,2,3,4)$ and $y = (7,4,4,-1)$ then

$$\text{distance between } x \text{ and } y = \sqrt{36 + 4 + 1 + 25} = \sqrt{66}$$

physical interpretation in R^2 (and similarly in R^3)

The 2-tuple (3,1) can be pictured as an arrow and as a point (Fig 2). If you do picture it as an arrow then it is be most useful to draw the arrow starting at the origin in which case the point with coords (3,1) is the head of the arrow.

The norm of a vector is the length of the arrow.

The arrow $-2u$ points in the direction opposite to u and is twice as long (Fig 3).

The arrow 3u points in the same direction as u and is 3 time as long (Fig 3).

There are triangle and parallelogram rules for picturing arrows u+v (Figs 4,5) and u−v (Fig 5).

The triangle inequality (property (4) of norms) says that the sum of two sides of a triangle is greater than the third side.

The arrow inclined at angle θ and with length r has components $(r\cos\theta, r\sin\theta)$ (Fig 6). In particular, the unit vector inclined at angle θ is $(\cos\theta, \sin\theta)$.

Two arrows u and v are perpendicular if $u \cdot v = 0$.

The arrow u_{unit} points like arrow u and has length 1.

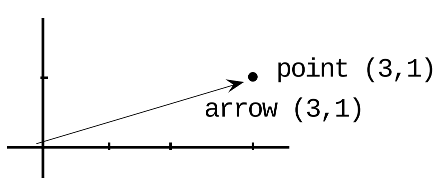


FIG 2

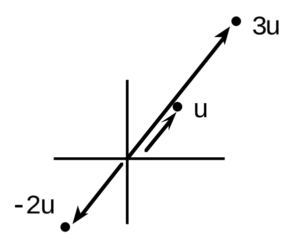


FIG 3

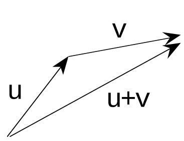


FIG 4

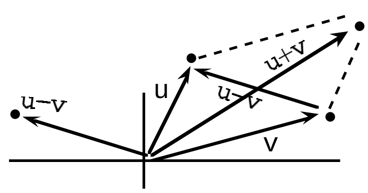


FIG 5

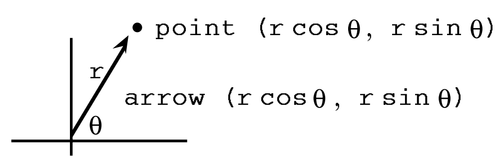


FIG 6

vector projection in R^2 and R^3

The vector projection (Figs 7a and 7b) of arrow u onto arrow v is $\dfrac{v \cdot u}{v \cdot v}\, v$.

The same formula works whether the angle between u and v is acute or an obtuse.

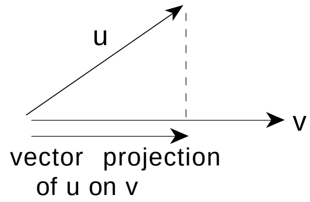


FIG 7a

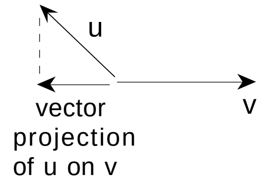


FIG 7b

footnote In case you didn't see vector projections in calculus here is why the formula works in both Fig 7a and in Fig 7b.

Remember from calculus that in Figs 8a and 8b below, $\cos\theta = \dfrac{u\cdot v}{\|u\|\ \|v\|}$.

In Fig 8a,

$$\begin{aligned} k_1 &= \|u\| \cos\theta \qquad \text{(right triangle trig)} \\ &= \|u\| \ \frac{u\cdot v}{\|u\|\ \|v\|} \\ &= \frac{u\cdot v}{\|v\|} \end{aligned}$$

The vector projection in Fig 7a points like v and has length k_1 so

$$\begin{aligned} \text{vector projection in Fig 7a } &= k_1\, v_{\text{unit}} \\ &= \frac{u\cdot v}{\|u\|\|v\|}\ \frac{v}{\|v\|} \\ &= \frac{u\cdot v}{\|v\|^2}\ v \\ &= \frac{u\cdot v}{v\cdot v}\ v \end{aligned}$$

In Fig 8b,

$$\begin{aligned} k_2 &= \|u\| \cos(\pi-\theta) \qquad \text{(right triangle trig)} \\ &= \|u\|\ (-\cos\theta) \qquad \text{(more trig)} \\ &= -\ \|u\|\ \frac{u\cdot v}{\|u\|\|v\|} \\ &= -\ \frac{u\cdot v}{\|v\|} \qquad \text{Note that this is positive since } \theta \text{ is obtuse in Fig 8b so } u\cdot v \text{ is neg} \end{aligned}$$

The vector projection Fig 7b points like $-v$ and has length k_2 so

$$\begin{aligned} \text{vector projection in Fig 7b} &= k_2\ (-v_{\text{unit}}) \\ &= -\ \frac{u\cdot v}{\|u\|\|v\|}\ \frac{-v}{\|v\|} \\ &= \frac{u\cdot v}{v\cdot v}\ v \qquad \text{(minuses cancel out)} \end{aligned}$$

Same formula in Figs 7a and 7b.

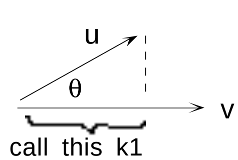


FIG 8a

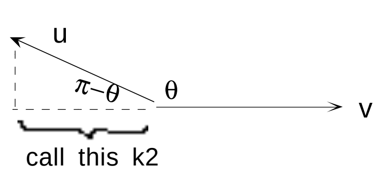


FIG 8b

connection between dotting vectors and multiplying matrices

If $u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $v = \begin{bmatrix} 4 \\ 6 \\ 5 \end{bmatrix}$ (vectors written as cols) then $u \cdot v = 4 + 12 + 15 = 31$.

This procedure is like doing the matrix multiplication

$$[1 \quad 2 \quad 3] \begin{bmatrix} 4 \\ 6 \\ 5 \end{bmatrix} \qquad \text{(row matrix times col matrix)}$$

In general, for column vectors u,v

(*) $\boxed{\text{the dot product } u \cdot v = \text{matrix product } u^T v}$

In particular

$$\boxed{u^T u = u \cdot u = \|u\|^2}$$

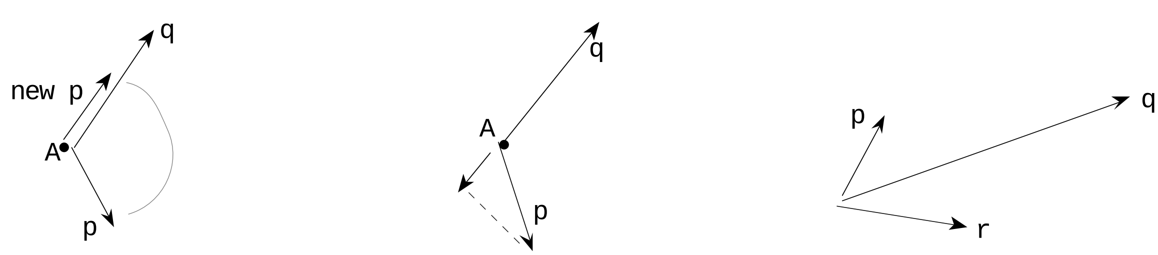
PROBLEMS FOR SECTION 2.1

1. Let $u = (2,3,-4,5)$ and $v = (3,-1,2,-6)$. Find

$$u + v, \quad u \cdot v, \quad \|u\|, \quad |u \cdot v|, \quad u_{\text{unit}}, \quad \frac{v \cdot u}{v \cdot v}\, v$$

2. Let $\vec{p} = \vec{i} - \vec{j}$ and $\vec{q} = 2\vec{i} + 6\vec{j}$. The diagrams show p and q drawn with common initial point A.
(a) Suppose p is rotated around so that it still has its tail at A but now it lies on top of q. Find (easily) the components of the new p.
(b) Suppose p is projected onto q. Find the components of the projection arrow.

3. The diagram shows vectors p,q,r. Express r roughly as a combination of p and q; i.e., estimate a and b so that $r = ap + bq$.



Problem 2(a) Problem 2(b) Problem 3

4. Show that $v - \dfrac{u \cdot v}{\|u\|^2}\, u$ is orthogonal to u

(a) in R^2 where you can do it just by drawing a picture

(b) in R^n in general

5. If $u \cdot v = 6$, $\|u\| = 3$, $\|v\| = 7$ find

(a) $\|-2u\|$ (b) $(u + v) \cdot (u - v)$ (c) $\|u + v\|$ (d) $(u + 3v) \cdot u$

6. This section claimed that $\frac{u}{\|u\|}$ is a unit vector. Use properties of the norm to prove that it is.

7. The dot product property $(ku) \cdot v = k(u \cdot v)$ has three kinds of multiplication in the one equation. Where are they?

8. Show that in R^n, $\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2 = 4(\vec{x} \cdot \vec{y})$.

9. Look at this statement: If $u \cdot v = 0$ then $\|u + v\| = \|u - v\|$.
(a) Draw pictures and use geometry to show that it is true in R^2.
(b) Use vector algebra to show that it's true in general in R^n.

10. Look at this statement: If $\|u\| = \|v\|$ then $u + v$ and $u - v$ are orthogonal.
(a) Draw pictures and use geometry to show that it is true in R^2.
(b) Use vector algebra to show that it's true in general in R^n.

11. It should be clear in R^2 that if u and v are perpendicular then u_{unit} and v_{unit} (which point the same way as u and v respectively) are also perpendicular. Show that it's true in R^n.

12. Consider the statement

(*) $$\|\vec{u} + \vec{v}\| = \|\vec{u}\| + \|\vec{v}\|$$

(a) Show that it is not always true in R^n.
To *prove* a general statement you have to give a general argument. But to *disprove* a general statement, as in this problem, you should find one specific counterexample.
(b) Draw pictures to show that it is true in R^2 if v is a positive multiple of u, and not true otherwise.
(c) Use vector algebra to show that it is true in R^n if v is a positive multiple of u.

SECTION 2.2 INDEPENDENT AND DEPENDENT SETS OF VECTORS

linear combinations of vectors

An expression like

$$4\vec{u} - 9\vec{v} + 8\vec{w} - \pi\vec{p}$$

is called a *(linear) combination* of u,v,w,p. Other combinations of u,v,w,p are

u (i.e., 1u + 0v + 0w + 0p)
6u
3w +2p
$3u - 2v + \sqrt{5}\,w + 6p$
2v + w − p
etc.

definition of independence and dependence

(1)

A set of vectors is called *dependent* if some vector in the set is a combination of other vectors in the set.
A set of vectors is called *independent* if no vector in the set is a combination of other vectors in the set.

In particular, a set of *two* vectors is dependent if one is a multiple of the other and is independent if neither is a multiple of the other.

For example, let

u = (1,0,0,0)
v = (0,1,0,0)
(*) w = (2,3,0,0)
x = (2,0,4,5)

Then

u,v are independent since neither vector is a multiple of the other;
u,v,x are independent;
u,v,w are dependent since w = 2u + 3v;
u,v,w,x are dependent since w = 2u + 3v.

Note that u,v,w,x in are dependent even though x is *not* a combination of the others. A dependent set may contain innocent vectors which are not combinations of the others. But as long as there's at least one vector that *is* a combination of others, the whole set is called dependent.

warning

In (*), w is a combination of u and v so the vectors u,v,w,x *collectively* are dependent but don't refer to w as a "dependent vector". Only a *set* of vectors can be called dependent, not individual vectors.

In (*), x is not a combination of u,v,w but don't call x an "independent vector". Only a *set* of vectors can be called independent, not individual vectors.

sets containing the zero vector

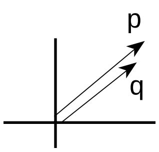
Any set containing $\vec{0}$ is a dependent set. For example, $\vec{0}$, u, v, w are dependent (no matter what u,v,w are) since $\vec{0}$ = 0u + 0v + 0w.

physical interpretation in R^2 and R^3

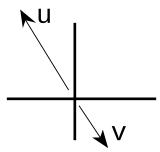
Imagine vectors as arrows with tail at the origin $\vec{0}$.

In R^2, two vectors are dependent if the arrows point in the same or in opposite directions (Figs 1 and 2). Otherwise they are independent (Fig 3).

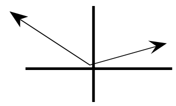
Fig 4 shows 3 vectors in R^2. You can see from the dashed parallelogram that $v = 2u + \frac{1}{2}w$ (approximately) so the vectors are dependent.



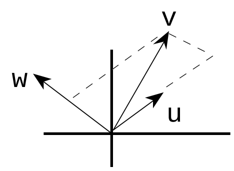
dep vectors
FIG 1



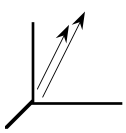
dep vectors
FIG 2



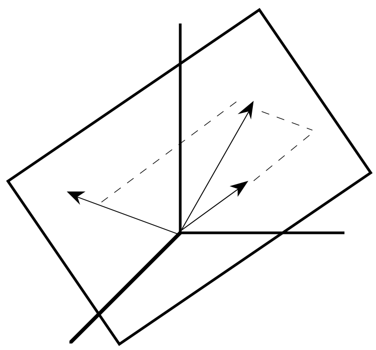
independent vectors
FIG 3



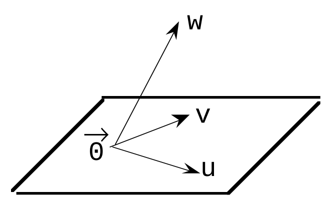
dep vectors
FIG 4



dep
FIG 5



dep
FIG 6



ind
FIG 7

In R^3, two vectors are dependent if the arrows point in the same or in opposite directions (Fig 5). Otherwise they are independent.

Three vectors in R^3 are dependent if the arrows are coplanar (Fig 6). Otherwise they are ind (Fig 7).

sets containing too many vectors to be independent

In R^2, a set of 3 or more vectors will always be dependent (as in Fig 4).
In R^3, a set of 4 or more vectors will always be dependent.
In general:

(2) $\boxed{\text{A set of } \textit{more} \text{ than n vectors in } R^n \text{ must be dependent}}$

A set of *n or fewer* vectors in R^n may be ind or dep, depending on the particular vectors.
(Proof of (2) coming later in §3.1.)

another test for independence

Start with say the vectors u,v,w,p.

To use the definition in (1) to show they are independent you have to show that u is not a comb of v,w,p and that v is not a comb of u,w,p and that w is not a comb of u,v,p and that p is not a comb of u,v,w.

Here's how to do it in one shot.

Try to solve the vector equation

(**) $$au + bv + cw + dp = \vec{0}$$

for the scalars a,b,c,d. One solution is a = 0, b = 0, c = 0, d = 0. This is called

the trivial solution.
See if you can do it some *other* way, i.e., with a, b, c, d not all 0.
If you *can't* do it some other way, i.e, if the *only* solution for a, b, c, d is the trivial solution then the vectors u,v,w,p are independent.
If you *can* do it some non-trivial way then the vectors are dependent.
Here's an illustration of why this works. Suppose the equation in (**) has the non-trivial solution $a=3$, $b=5$, $c=2$, $d=0$, so that $3u + 5v + 2w + 0p = \vec{0}$. Then the vectors u,v,w,p are dependent because you can write say u as a combination of v,w,p, namely $u = -\frac{5}{3}v - \frac{2}{3}w$.

(3)

> To decide if u,v,w are ind or dep solve the equation $au + bv + cw + dp = \vec{0}$ for a,b,c,d.
> The equation always has the trivial solution a=0, b=0, c=0, d=0.
>
> The vectors are ind iff the equ has *only* the trivial sol a=0, b=0, c=0, d=0.
>
> The vectors are dep iff the equ has *at least one other* solution for a,b,c,d in addition to the trivial solution.

example 1

Let

$$\vec{v}_1 = (1,0,-2,2), \quad \vec{v}_2 = (0,0,1,1), \quad \vec{v}_3 = (2,0,-1,7), \quad \vec{v}_4 = (0,1,0,0)$$

(a) Are the vectors dep or ind.
(b) If they are dep find a relation among them, i.e., write one of them as a combination of others.

solution (a) Solve the vector equation

$$a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3 + d\vec{v}_4 = \vec{0}$$

This is

$$a(1,0,-2,2) + b(0,0,1,1) + c(2,0,-1,7) + d(0,1,0,0) = (0,0,0,0)$$

$$(a,0,-2a,2a) + (0,0,b,b) + (2c,0,-c,7c) + (0,d,0,0) = (0,0,0,0)$$

$$(a+2c,\ d,\ -2a+b-c,\ 2a+b+7c)\ \ (0,0,0,0)$$

which amounts to the system of equations

$$\begin{aligned} 1a + 0b + 2c + 0d &= 0 \\ 0a + 0b + 0c + 1d &= 0 \\ -2a + 1b - 1c + 0d &= 0 \\ 2a + 1b + 7c + 0d &= 0 \end{aligned}$$

Here's the solution found by Mathematica.

```
In[3]
v1 = {1,0,-2,2}; v2 ={0,0,1,1}; v3 = {2,0,-1,7}; v4 = {0,1,0,0};
In[4]
Solve[a v1 + b v2 + c v3 + d v4 == {0,0,0,0}]
Out[4]
{{a -> -2 c, b -> -3 c, d -> 0}}
```

The solution is d = 0, c = anything, b = −3c, a = −2c.
You can get solutions by choosing any c you like.
One solution (choose c=0) is the trivial sol d = 0, c = 0, b = 0, a = 0.

Another solution (choose c=1) is d = 0, c = 1, b = −3, a = −2.
Another solution (choose c=2) is d = 0, c = 2, b = −3, a = −4 etc.
There is a not-all-zero solution (lots of them), so the vectors are dep.
(b) Here's how to actually write one of the vectors as a combination of the others. Using say the solution d = 0, c = 1, b = −3, a = −2 we have

$$-2\,v_1 - 3\,v_2 + v_3 = \vec{0}\ .$$

So $v_3 = 2v_1 + 3v_2$.

example 2

Let

$$\vec{u}_1 = (0,1,2,4),\quad \vec{u}_2 = (3,-1,-1,0),\quad \vec{u}_3 = (2,-3,-2,8),\quad \vec{u}_4 = (1,0,2,2)$$

And here's the solution (by Mathematica) to the system of equations

$$a\,\vec{u}_1 + b\,\vec{u}_2 + c\,\vec{u}_3 + d\,\vec{u}_4 = \vec{0}$$

```
In[1]
u1 = {0,1,2,4}; u2 ={3,-1,-1,0}; u3 = {2,-3,-2,8}; u4 = {1,0,2,2};
In[2]
Solve[a u1 + b u2 + c u3 + d u4 =={0,0,0,0}]
Out[2]
{{a -> 0, b -> 0, c -> 0, d -> 0}}
```

Since the only sol is a=0, b=0, c=0, d=0, the vectors u_1, u_2, u_3, u_4 are independent.

example 3

Suppose u,v,w are independent. Show that u, v, 3u + 2w are independent.

solution Try to solve

(***) $$au + bv + c(3u + 2w) = \vec{0}$$

for a, b, c. Rearrange the equation:

$$(a+3c)u + bv + 2cw = \vec{0}$$

The vectors u,v,w are ind by hypothesis so by (3),

$$a+3c = 0,\ b = 0,\ 2c = 0$$

But then c = 0, b = 0, a = 0.
So the only solution to (***) is a = b = c = 0.
That makes u, v, 3u + 2w ind, by (3) again.

and still another test for independence

This is a refinement of (1). It's a little shorter than testing to see that no vector is a combination of *any* of the other vectors.

(4) If the first vector on a list is not $\vec{0}$ and after that no vector on the list is a combination of the *preceding* vectors then the vectors are independent.

In other words, to test u,v,w,p: if u is not 0 and v is not a multiple of u, and w is not a combination of u and v, and p is not a combination of u,v,w then the vectors are independent.

Or you could list the vectors in the order v,p,w,u. Then if v is not $\vec{0}$ and p is not a multiple of v, and w is not a combination of v and p, and u is not a combination of v,p,w then the vectors are ind.

You get to choose in what order you list the vectors for this test.

example 4

Let u = (6,7,8,9), v = (1,0,0,0), p = (0,1,0,0), r = (0,0,1,0).

To test for independence, look at the vectors reading from right to left (because it's that's a more convenient order).

r is not $\vec{0}$.
p is not a multiple of r
v is not a combination of p and r
(because the first component of v is nonzlero but the first components of p and r are both 0 so there is no way that v can be ap + br)

u is not a combination of of r,p,v
(because the fourth component of u is nonzero but the fourth components of r,p and v are all 0)

So the vectors u,v,p,r are ind.

orthogonality and independence

(5) $\boxed{\text{Any set of nonzero } \textit{orthogonal} \text{ vectors in } R^n \text{ must be independent}}$

(But this is only a one-way test: if a set of vectors is *not* orthogonal, you *can't* conclude that they are dependent.)
(See problem 6 for the proof of (5).

unique representation rule

Suppose $\vec{x}$ can be written as a combination of u,v,w, (e.g., $\vec{x}$ = 8u + 9v + 10w).

(a) If u,v,w are dependent then x can be written in still more ways as a combination of u,v,w.
(b) If u,v,w are independent then x can be written in just this one way as a combination of u,v,w.

In other words, assuming that x can be written at all as a combination of u,v,w then unique representation in terms of u,v,w goes with u,v,w independent; and non-unique representation in terms of u,v,w goes with u,v,w dependent.

inadequate proof of (a)

Suppose x = 8u + 9v + 10w.
Suppose u,v,w are dependent. Then one of them is a combination of others, say v=3u.
Then x = 8u + 9(3u) + 10w = 35u + 10w. So the original representation of x as 8u+9v+10w was not unique.

proof of (b)

Suppose x = 8u + 9v + 10w.
Suppose u,v,w are independent.
I want to show that x can't be written in a *second* way as a combination of u,v,w.
Here's a proof by contradiction.
Suppose there were a second way to write x as a combination of u,v,w, say x = 12w.
Then 8u + 9v + 10w = 12w so $u = -\frac{9}{8}v + \frac{2}{8}w$ which contradicts the fact that u,v,w are ind.
So there can't be a second way to write x as a combination of u,v,w.

mathematical catechism (you should know the answers to these questions)

question 1 What does it mean to say that vectors u,v,w,p are independent.

first answer It means no one of them is a combination of any of the others.

second answer It means the only solution to au+bv+cw+dp = $\vec{0}$ is a=0, b=0, c=0, d=0.

question 2 What does it mean to say that vectors u,v,w,p are dependent.
first answer It means that one of them is a combination of the others.

second answer It means that the vector equation au+bv+cw+dp = $\vec{0}$ has a solution for a,b,c,d besides the trivial solution a=0, b=0, c=0, d=0.

PROBLEMS FOR SECTION 2.2

1. True or False. If True explain why. If False find a counterexample.
(a) If u, v, w, x are ind then u,v,w are ind.
(b) If u,v,w are ind then u, v, w, x are ind.
(c) If u,v,w,s are dep then u,v,w are dep.
(d) If u,v,w are dep then u,v,w,x are dep.
(e) If u,v are ind and u,w are ind and v,w are ind then u,v,w are ind.

2. Let u = (2,3,4,5,6), v = (1,0,1,0,1), w = (5, 6, 9, 10, 13), p = (−1, 3, 1, 5, 3).
Here's the solution (by Mathematica) to the vector equation au + bv + cw + dp = $\vec{0}$.

```
In[1
u = {2,3,4,5,6}; v = {1,0,1,0,1}; w = {5,6,9,10,13}; p = {-1,3,1,5,3};
Solve[a u + b v + c w + d p == {0,0,0,0,0}]
Out[1]
{{a -> -2 c - d, b -> -c + 3 d}}
```

(a) The vector equation really a system of how many equations in how many unknowns. Write out the system to make sure you understand it.
(b) Are u,v,w,p independent or dependent.
If they are dependent, write one of them as a combination of others.

3. Suppose I change some of the components of the vectors u,v,w,p in problem 2 so that the vectors are now independent. What would the output of the Mathematica program be in that case.

4. If u,v,w are independent show that the vectors u + v, v + w, w + u are also ind.

5. Let u = (1,0,0,0), v = (1,1,0,0), w = (1,1,1,0), p = (1,1,1,1).
Test for independence using
(a) (3)
(b) (4)

6. Suppose u,v,w are in R^n and A is an n × n matrix. If you think of u,v,w as column matrices then you can left multiply them by A to get the new vectors (actually column matrices) Au, Av, Aw.
Show that if u,v,w are independent and A is invertible then Au, Av, Aw are also ind.

7. Prove (5) by showing that if u,v,w are nonzero orthogonal vectors in R^n then they are independent. Suggestion: Try to solve au + bv + cw = $\vec{0}$ by dotting both sides with u.

8. Is it possible to have 10 orthogonal nonzero vectors in
(a) R^3 (b) R^4 (c) R^9 (d) R^{10}

SECTION 2.3 BASES

coordinate systems and bases in R^2

Fig 1 shows an X,Y coordinate system superimposed on the usual x,y system. Point p has coordinates x = 4, y = 2 and also has coords X = 5, Y = −1.

In a coordinate system in R^2, the vectors from the origin to the unit points on the axes are called *basis vectors*.

For the standard x,y coordinate system in Fig 1, the basis vectors are

$$\vec{i} = (1,0) \quad \text{and} \quad \vec{j} = (0,1).$$

For the X,Y coordinate system in Fig 1, the basis vectors are

$$\vec{u} = (1,1) = \vec{i} + \vec{j} \quad \text{and} \quad \vec{v} = (1,3) = \vec{i} + 3\vec{j}$$

Since $\|u\| = \sqrt{2}$, the scale on the X-axis is $\sqrt{2}$ times the scale in the x,y system.
Since $\|v\| = \sqrt{10}$, the scale on the Y-axis is $\sqrt{10}$ times the scale in the x,y system.

When you use the standard coordinate system you are expressing vectors as combinations of the basis vectors i and j. Writing p = (4,2) is the same as writing p = 4i + 2j

When you use the new X,Y coord system you are expressing vectors as combinations of the basis vectors u and v. Saying that p has coords X = 5, Y = −1 is the same as saying p = 5u − v.

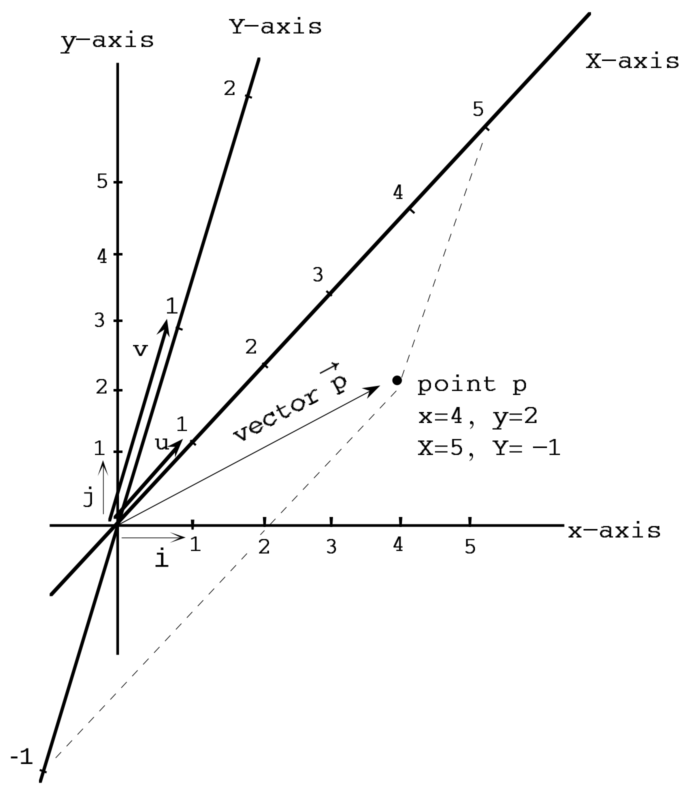


FIG 1

spanning vectors

The vectors u,v,w, .. *span* R^n if every vector in R^n can be written as a combination of u,v,w, ...

For example, if u = (2,0), v = (0,3), w = (2,3) then u,v,w span R^2 because every vector in R^2 can be written in terms of them. In fact, u,v by themselves span R^2. (And so do u,w by themselves. And so do v,w by themselves.)

definition of a basis for R^n

(1)

A basis for R^n is a bunch of vectors $\vec{u}$, $\vec{v}$, $\vec{w}$, ... in R^n such that every vector in R^n can be written in exactly one way in terms of them.

In other words, a basis for R^n is a bunch of vectors $\vec{u}$, $\vec{v}$, $\vec{w}$, ... in R^n with these two properties.

(a) $\vec{u}$, $\vec{v}$, $\vec{w}$, ... span R^n.

(b) When you write a vector in terms of $\vec{u}$, $\vec{v}$, $\vec{w}$, ... it can't be done in more than one way.

the n independent vectors rule (an equivalent characterization of a basis for R^n)

(2a) If a bunch of vectors is a basis for R^n (i.e., if (1a) and (1b) hold) then there must be n vectors in the bunch and the vectors must be independent.

(2b) Conversely, if you have n independent vectors in R^n then they are a basis (i.e., (1a) and (1b) hold).

R^n is called n-dimensional because it takes n vectors to make a basis for R^n.

partial proof of (2a)

The fact that the vectors in a basis must be independent follows from the unique representation rule in the preceding section.

I'm leaving out the proof that there must be n of them.

proof of (2b) Coming up in Section 3.1.

coordinates of a vector w.r.t. a basis

If $\overrightarrow{u_1}$, $\overrightarrow{u_2}$, ..., $\overrightarrow{u_n}$ is a basis for R^n and $p = a_1 \overrightarrow{u_1} + a_2 \overrightarrow{u_2} + \ldots + a_n \overrightarrow{u_n}$ then the scalars a_1, a_2, ..., a_n are called the *coordinates* or *components* of p with respect to the basis. What (1a) and (1b) together say is that every point has exactly one set of coordinates w.r.t. a basis.

the standard basis for R^n

The standard basis for R^2 is

$$i = (1,0), \qquad j = (0,1)$$

The vectors u,v in Fig 1 are another basis for R^2.

The standard basis for R^3 is

$$i = (1,0,0), \quad j = (0,1,0), \quad k = (0,0,1)$$

The standard basis for R^4 is

$$i = (1,0,0,0), \quad j = (0,1,0,0), \quad k = (0,0,1,0), \quad \ell = (0,0,0,1)$$

For example, if u = (2,3,4,5) then

$$u = 2i + 3j + 4k + 5\ell$$

The coords of u with respect to the standard basis are 2,3,4,5. These are called the *natural* coords of u.

In general, the standard basis in R^n is

$$\vec{e}_1 = (1,0,0, \ldots,0), \quad \vec{e}_2 = (0,1,0,0, \ldots,0), \ldots, \quad \vec{e}_n = (0,0, \ldots, 0,1)$$

example 1

Another basis for R^4 is

$$p = (2,0,0,0),\ q = (1,1,0,0),\ r = (0,0,1,0),\ s = (0,0,1,1)$$

because there are 4 vectors and I checked that they are independent.
If u = (2,3,4,5), to express u in terms of p,q,r,s we want to find a,b,c,d so that

$$(2,3,4,5) = a(2,0,0,0) + b(1,1,0,0) + c(0,0,1,0) + d(0,0,1,1).$$

This means solving the system

$$\begin{aligned} 2 &= 2a + 1b + 0c + 0d \\ 3 &= 0a + 1b + 0c + 0d \\ 4 &= 0a + 0b + 1c + 1d \\ 5 &= 0a + 0b + 0c + 1d \end{aligned}$$

There is just one solution, namely, d = 5, b = 3, a = −1/2, c = −1 so

$$u = -\frac{1}{2}p + 3q - r + 5s$$

The coordinates of u w.r.t. the basis p,q,r,s are −1/2 , 3, −1, 5.

unofficial point of view

You can imagine that a set of basis vectors in R^n determines axes for a new coordinate system (as in Fig 1).

If u is one of the basis vectors and $\|u\| = 3$ then imagine that the scale on the axis determined by u is 3 times the scale in the standard $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ system.

If a basis is normalized (i.e., the basis vectors all have norm 1) then the scale on each new axis is the same as the scale in the standard system.

If a basis for R^n is orthogonal (like $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ but unlike p,q,r,s in (2)), then the corresponding axes (which exist only in your imagination) are perpendicular.

terminology

The plural of basis is bases (as in "R^n has many bases").
The vectors in a basis are called basis vectors.

example 2

Fig 2 shows the coordinate system with basis

$$u = -3i + j, \quad v = 3i + 2j$$

(a) The point q has coords x = −3, y = 4 in the standard coord system. Find it's

coordinates w.r.t. u and v.
(b) How do the scales in the new coord system compare with the old scale.

solution (a) We want to find X and Y so that q = Xu + Yv, i.e., we want to solve

$$(-3,4) = X(-3,1) + Y(3,2).$$

This is the system

$$\begin{aligned} -3 &= -3X + 3Y \\ 4 &= X + 2Y. \end{aligned}$$

The solution is X = 2, Y = 1 so q has coords X = 2, Y = 1 in the new coord system.
(b) $\|u\| = \sqrt{10}$ so the scale on the X-axis is $\sqrt{10}$ times the scale in the old x,y system.
$\|v\| = \sqrt{13}$ so the scale on the Y-axis is $\sqrt{13}$ times the old scale.

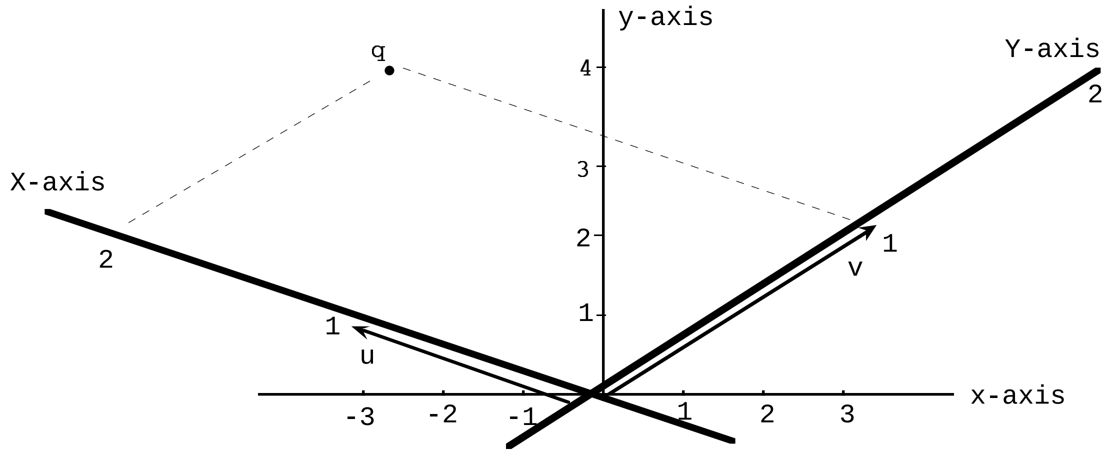


FIG 2

warning

1. In the standard coordinate system in R^2 you use the letter x to name the x-axis and you also use it as the name for the first coordinate. But you don't use it to name the basis vector which points along the x-axis. The standard name for that basis vector is i, not x.

Similarly, when you switch to another coordinate system in R^2, if you use the letter X as the name of a first coordinate then you shouldn't also use X to name one of the new basis vectors. Use a name like $\vec{u}$ or $\vec{v}$ for the basis vector, not X.

2. To say that the X scale is twice the x scale you can write X-scale = 2 x-scale. But *don't* write X = 2x. The equation X = 2x says that a point's X coord is twice its x coord which is not what you mean.

3. The notation $\vec{p} = (2,4,6)$ for a vector in R^3 means $\vec{p} = 2\vec{i} + 4\vec{j} + 6\vec{k}$. If you use basis $\vec{u},\vec{v},\vec{w}$ and $\vec{p} = -\vec{u} + 6\vec{v} + 5\vec{w}$ don't write $\vec{p} = (-1,6,5)$ since people will think you mean $-\vec{i} + 6\vec{j} + 5\vec{k}$. Invent some new notation such as $((-1,6,5))$ or $(-1,6,5)_{uvw}$ or just leave it $-\vec{u} + 6\vec{v} + 5\vec{w}$.

4. When you draw the standard coord system in R^2, you put the label x-axis on the *positive* half of the axis. Similarly, when you draw a new coord system, put the label X-axis on the *positive* half of the axes, and similarly for the Y-axis.

the n orthog rule (corollary of (2b))

(3) $\boxed{\text{A set of n nonzero orthogonal vectors in } R^n \text{ is a basis for } R^n.}$

This is only a one-way rule. It says that if you have n nonzero orthog vectors in R^n then they are a basis. It does not say that if you have a basis for R^n then the vectors must be orthogonal.

proof of (3)

Nonzero orthogonal vectors are independent (from the preceding section).

So a set of n nonzero orthog vectors in R^n is a set of n *independent* vectors in R^n.

So they are a basis for R^n by (2b).

summary: to show that a bunch of vectors is a basis for R^n

method 1 Show that the vectors span R^n uniquely (i.e., show that (1a) and (1b) hold).
method 2 (*this is probably the most practical*) Check that there are n of them and show that they are independent.
method 3 (a special case of the preceding) If there are n of them and they are nonzero orthogonal then they are a basis.

What you do with basis vectors for R^n (i.e., what is their role in life)

You express vectors (uniquely) in terms of them.

example 3

Let u = (1,2,3,4), v = (0,1,1,1), w = (0,0,1,1), p = (0,0,0,1).

(a) Show that they are a basis for R^4.
(b) Let x = (3,5,10,14). Find the coordinates of x w.r.t. basis u,v,w,p.

solution (a) First I'll show that u,v,w,p are independent using (4) from the preceding section.

Look at the vectors on the list and read from right to left (for convenience).

You can see that $p \ne \vec{0}$.
And w is not a multiple of p.
And v is not a combination of w and p because the second component of v is non-zero but the second components of p and w are 0.
And u is not a combination of the preceding vectors p,w,v because the first component of u is non-zero but the first components of p,v,w are 0. This makes p,w,v,u independent.

So u,v,w,p are four independent vectors in R^4. By (2b) they are a basis for R^4.

(b) Maybe you can see by inspection that

$$x = 3u - v + 2w + p$$

If you're not good at inspecting then just solve the vector equation

$$au + bv + cw + dp = x$$

which is the system of equations

$$\begin{aligned} a \qquad\qquad &= 3 \\ 2a + b \qquad &= 5 \\ 3a + b + c &= 10 \end{aligned}$$

$4a + b + c + d = 14$

The solution is a = 3, b = −1, c = 2, d = 1
So the coords of x w.r.t. basis u,v,w,p are 3,−1,2,1.

example 4

Let

u = (1,2,3,4); v = (5,6,3,2); w = (1,1,1,1); p = (1,0,1,0); x = (2,2,1,1);

(a) Show that u,v,w,p are a basis for R^4.
(b) Find the coordinates of x w.r.t. the basis u,v,w,p.

solution (a) Can't do much with these vectors by inspection but using a computer, here's the test that u,v,w,p are independent.

```
u = {1,2,3,4}; v = {5,6,3,2}; w = {1,1,1,1}; p = {1,0,1,0};
Solve[a u + b v + c w + d p == {0,0,0,0}]
{{a -> 0, b -> 0, c -> 0, d -> 0}}
```

Since the only solution to au + bv + cw + dp = $\vec{0}$ is a = b = c = d = 0, the vectors u,v,w,p are independent. Since there are four of them, they are a basis for R^4.
(b) To get the coords of x w.r.t. the basis u,v,w,p, solve au + bv + cw + dp = x

```
x = {2,2,1,1};
Solve[a u + b v + c w + d p == x]
              1                           1
{{a -> -(-), b -> 0, c -> 3, d -> -(-)}}
              2                           2
```

The solution is a = −1/2, b = 0, c = 3, d = − 1/2.
So the coords of x w.r.t. the basis u,v,w,p are − 1/2, 0, 3, −1/2.

example 5 (a non-basis)

Let u = (1,5,4,9), v = (2,1,2,3), w = (5,4,1,2).

The vectors u,v,w are not a basis for R^4 because there aren't enough of them. Some vectors can be written in terms of u,v,w but not every vector can. Here's what happens when you use Mathematica to try to write x = (1,2,3,4) in terms of u,v,w

```
In[1]
u = {1,5,4,9}; v = {2,1,2,3}; w = {5,4,1,2}; x = {1,2,3,4};
Solve[a u + b v + c w == x]
Out[1]
{}
```

There are no solutions for a,b,c. So x can't be written in terms of u,v,w.

example 6 (a non-basis)

Let u = (1,0,0,0), v = (0,1,0,0), w = (0,0,1,0), p = (0,0,0,1), q = (1,1,1,1).

The vectors u,v,w,p,q are not a basis for R^4 because there are too many of them. The vectors do happen to span R^4 (the first four are a basis for R^4) but they are not independent. If x = (1,2,3,4), here's what happens when you use Mathematica to try to write x in terms of u,v,w,p,q.

```
In[2]
u = {1,0,0,0}; v={0,1,0,0}; w={0,0,1,0};  p={0,0,0,1}; q={1,1,1,1};
x = {1,2,3,4};
Solve[a1 u + a2 v + a3 w + a4 p + a5 q == x]
Out[2]
{{a1 -> 1 - a5, a2 -> 2 - a5, a3 -> 3 - a5, a4 -> 4 - a5}}
```

There are infinitely many possibilities.
For instance you can choose a5 = 0, in which case a4 = 4, a3 = 3, a2 = 2, a1 = 1:

$$x = u + 2v + 3w + 4p + 0q$$

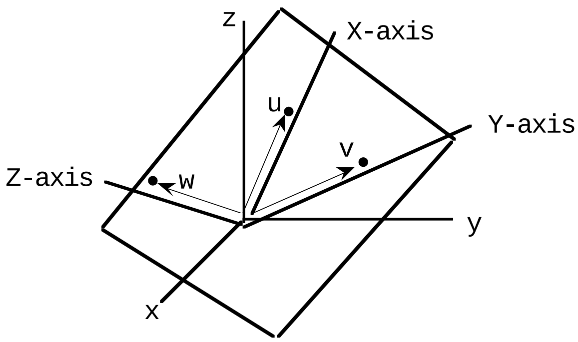
Or you can choose $a5 = 1$ in which case $a4 = 3$, $a3 = 2$, $a2 = 1$, $a1 = 0$:

$$x = 0u + v + 2w + 3p + q$$

And so on. The vectors u,v,w,p,q span R^4 but are not a basis because vectors cannot be written *uniquely* in terms of u,v,w,p,q.

example 7 (a non-basis)

If you have the right number of vectors but they are *dependent* then they won't span. Fig 3 shows three dependent vectors u,v,w in R^3, all in the same plane through the origin. Points not in that plane can't be written in terms of u,v,w so u,v,w don't span R^3. There may be enough axes but they aren't "decent" axes.



The X,Y,Z system is a *non*-coord system for R^3 because the X,Y,Z axes are all in the same plane.
FIG 3

dots and norms in a new coordinate system

Look at the vector p in Fig 1. In the usual coordinate system,

$$p = 4i + 2j = (4,2)$$

$$\|p\| = \sqrt{16 + 4} = \sqrt{20} \quad \text{(the length of the arrow p)}.$$

Fig 4 shows why this calculation produces the length of p, namely because p is the hypotenuse of a right triangle with sides 4 and 2.

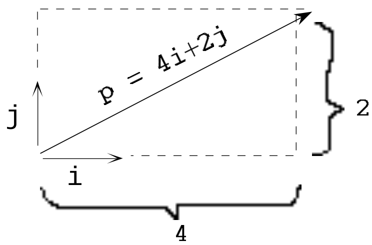
In the *new* coord system,

$$p = 5u - v = ((5,-1))$$

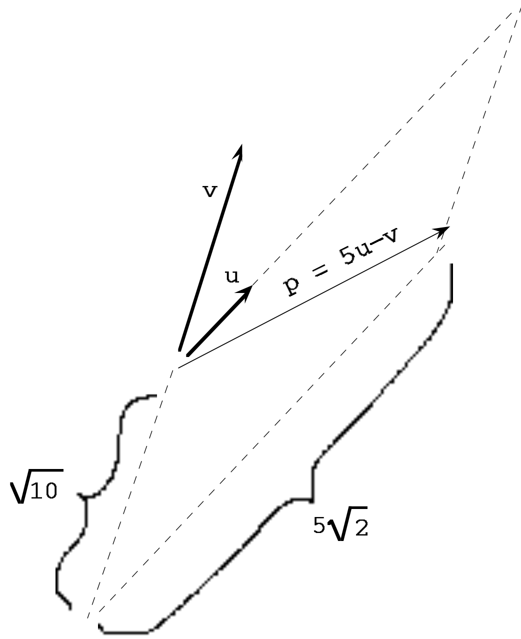
but

$$\|p\| \text{ is } \textit{not } \sqrt{25 + 1} \;.$$

You can see why not in Fig 5 where (unlike Fig 4) p is *not* the hypotenuse of a right triangle with sides 5 and 1. Instead, p is the third side of a non-right triangle with sides $\|5u\| = 5\sqrt{2}$ and $\|-v\| = \sqrt{10}$. So if you take $\sqrt{25 + 1}$ you won't get the length of p. It would have worked if u and v had been orthonormal.



$\|p\|$ is $\sqrt{16 + 4}$
FIG 4



$\|p\|$ is not $\sqrt{25 + 1}$
FIG 5

Here's the idea in general, for dots as well as for norms.

> Suppose $\vec{u_1}, \ldots, \vec{u_n}$ is an *orthonormal* basis for R^n. And suppose
>
> $$\vec{x} = X_1\vec{u_1} + \ldots + X_n\vec{u_n} = ((X_1, \ldots, X_n))$$
>
> $$\vec{y} = Y_1\vec{u_1} + \ldots + Y_n\vec{u_n} = ((Y_1, \ldots, Y_n))$$
>
> Then
>
> (6) $$\|x\| = \sqrt{X_1^2 + \ldots + X_n^2}$$
>
> (7) $$x \cdot y = X_1Y_1 + \ldots + X_nY_n$$
>
> In other words, the coordinates w.r.t. the basis $\vec{u_1}, \ldots, \vec{u_n}$ (in the double parentheses) can be maneuvered like ordinary $\vec{i}, \vec{j}$ coordinates to find dot products and norms.
>
> But if $\vec{u_1}, \ldots, \vec{u_n}$ are not orthonormal then (6) and (7) are not necessarily true.

For example suppose u,v is a basis for R^2 and

$$\vec{x} = 2u - 4v = ((2,-4))$$

$$\vec{y} = 9u + 7v = (((9,7))$$

Then it is *not* necessarily true that $\|x\| = \sqrt{4 + 16} = \sqrt{20}$ and it is not necessarily true that $x \cdot y = 18 - 28 = -10$. But it *would* be true if u,v were an orthonormal basis.

proof of (7)

Here's a proof for R^3. The R^n proof is similar.

Suppose $\vec{u_1}$, $\vec{u_2}$, $\vec{u_3}$ is an orthonormal basis for R^3 and

$$\vec{x} = X_1\vec{u_1} + X_2\vec{u_2} + X_3\vec{u_3} \quad \text{and} \quad \vec{y} = Y_1\vec{u_1} + Y_2\vec{u_2} + Y_3\vec{u_3}$$

Then

$$x \cdot y = (X_1\vec{u_1} + X_2\vec{u_2} + X_3\vec{u_3}) \cdot (Y_1\vec{u_1} + Y_2\vec{u_2} + Y_3\vec{u_3})$$

(8) $$= X_1Y_1\ u_1 \cdot u_1 + X_2Y_2\ u_2 \cdot u_2 + X_3Y_3\ u_3 \cdot u_3 + 2X_1Y_2\ u_1 \cdot u_2 + 2X_1Y_3\ u_1 \cdot u_3 + 2X_2Y_3\ u_2 \cdot u_3 \quad \text{by dot rules}$$

Since u_1, u_2, u_3 are orthonormal,

$$u_1 \cdot u_1 = \|u_1\|^2 = 1, \quad u_2 \cdot u_2 = 1, \quad u_3 \cdot u_3 = 1$$
$$u_1 \cdot u_2 = 0, \ u_1 \cdot u_3 = 0, \ u_2 \cdot u_3 = 0$$

and (8) becomes

$$x \cdot y = X_1Y_1 + X_2Y_2 + X_3Y_3 \qquad \text{QED}$$

mathematical catechism (you should know the answers to these questions)

question 1 What does it mean to say that u,v,w,r,... span R^{23}.

answer It means that every vector in R^{23} can be written as a combination of u,v,w,r,...

question 2 What does it mean (according to the definition) to say that u,v,w,r,... is a basis for R^{23}.

answer It means that every vector in R^{23} can be written in exactly one way in terms of u,v,w,r,...

question 3 If u,v,w,q is a basis for R^4, what does it mean to say that $\vec{p}$ has coordinates 2,3,9,−1 w.r.t.that basis.

answer It means that $p = 2u + 3v + 9w - q$.

question 4 What is R^4.

answer R^4 the set of all 4-tuples, i.e., the set of all (x_1, x_2, x_3, x_4)'s.

(Well, actually just 4-tuples of *real* numbers but for the time being we're only considering real numbers so I take that for granted.)

PROBLEMS FOR SECTION 2.3

1. Let $\vec{x} = (2,3,4)$.

(a) Let $u = (-1,0,0)$, $v = (0,-2,0)$, $w = (0,0,-1)$.

(i) Check that u,v,w is a basis for R^3.

(ii) Find the coordinates of x w.r.t. basis u,v,w.

(b) Let $p = (1,0,0)$, $q = (0,1,1)$, $r = (0,0,1)$

(i) Check that p,q,r is a basis for R^3.

(ii) Find the coordinates of x w.r.t. each basis.

2. Let $u = 5i + j$, $v = 5i - 5j$

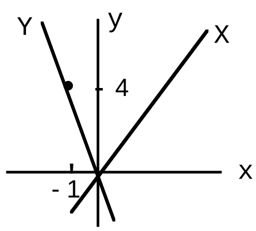
(i) How do you know that u,v is a basis for R^2.

(ii) Sketch the coord system with basis u,v (label and calibrate all the axes) .

3. Suppose the scale in the x,y system is the inch. A new X,Y system uses the same axes but has the foot as its unit. Find the new basis vectors.

4. The diagram shows a new coordinate system.

The X-axis isline y = 3x and the scale on the X-axis is the same as the scale in the old coord system.
The unit point on the Y-axis is point (−1,4).
Find the new basis vectors.



Problem 4

5. Suppose the scale in the x,y system is the foot. A new X,Y coord system has basis $u = 3j$ and $v = \frac{1}{2}i$.

(a) Describe the new axes and new scales.
(b) Find the connection between X,Y and x,y. In particular, express X and Y in terms of x and y.

6. (a) What's the difference between saying that a bunch of vectors spans R^3 and saying that a bunch of vectors is a basis for R^3.

(a) Find some vectors that span R^3 but are not a basis.
(b) Can you find some vectors that are a basis for R^3 but do not span R^3.

7. Let u = (1,0,0,0), v = (1,0,1,0), w = (0,0,0,1), p = (2,0,1,0).
(a) Let x = (3,,0,5,6). Write x as a combination of u,v,w,p in at least three ways,
(b) Let y = (3,4,5,6). Show that y can't be written as a combination of u,v,w,p.
(c) Are u,v,w,p a basis for R^4

8. I started with

$u = (\pi, \sqrt{13}, 4, 2)$
$v = (\cdot,\cdot,\cdot,\cdot)$ (I have specific coordinates in mind but you would't need to use them so I'm not bothering to write them down here)
$w = (\cdot,\cdot,\cdot,\cdot)$
$p = (\cdot,\cdot,\cdot,\cdot)$

And I picked $x = (2,0,-1,\sqrt{13})$.
Then I used Mathematica to solve the vector equation x = au + bv + cw + dp.
If possible, decide if u,v,w,p is a basis for R^4 if

(a) the solution came out to be

```
              1    c        1   c
{{a -> -(-) - -, b -> - - -, d -> -1}}
              2    2        2   2
```

(b) there is no solution.

9. Suppose u,v is a basis for R^2 and p = 2u+ 4v, q = 3u + 5v.
Find, if possible, the numerical value of $p\cdot q$ and $\|q\|$ if
(a) you have no further information about u,v
(b) u, v are orthonormal
(c) $\|u\| = 3$, $\|v\| = 4$, $u\cdot v = 5$

10. Suppose u,v,w is an orthogonal basis for R^3, $\|u\| = 2$, $\|v\| = 3$, $\|w\| = 1/4$ and $x = 4u + 5v + 9w$. (a) Write x in terms of the orthonormal basis u_{unit}, v_{unit}, w_{unit}.
(b) Find $\|x\|$.

11. I have a bunch of vectors in R^6. I didn't count them but they do span R^6 and they are independent. Are they a basis. Defend your answer.

SECTION 2.4 SOME BASIS CHANGING RULES

the basis changing matrix

Suppose $q = (5,6,7)$, i.e., $q = 5i + 6j + 7k$.
Let

$$u = (3,2,5),\ v = (1,6,9),\ w = (2,4,6)$$

be a new basis for R^3. (I checked that u,v,w they are independent so they are a basis.)

To get the new coords a,b,c of q w.r.t. basis u,v,w, solve

(1) $$q = au + bv + cw.$$

I want to see how the new coords a,b,c are related to the old coordinates 5,6,7.

The vector equation in (1) is

$$(5,6,7) = a(3,2,5) + b(1,6,9) + c(2,4,6)$$

This is the system of equations

$$\begin{aligned} 5 &= 3a + b + 2c \\ 6 &= 2a + 6b + 4c \\ 7 &= 5a + 9b + 6c \end{aligned}$$

which in matrix form is

(2) $$\underbrace{\begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}}_{\text{old coords}} = \underbrace{\begin{bmatrix} 3 & 1 & 2 \\ 2 & 6 & 4 \\ 5 & 9 & 6 \end{bmatrix}}_{\text{call this P}} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{\text{new coords}}$$

Note that the columns of the matrix P in (2) are the new basis vectors u,v,w.
In general:

(3)

Let $\vec{u},\vec{v},\vec{w}$ be a new basis for R^3 (same idea works in R^n).
The old basis is i,j,k.

Let $P = \begin{bmatrix} \vec{u} & \vec{v} & \vec{w} \end{bmatrix}$, the matrix whose cols are $\vec{u},\vec{v},\vec{w}$; in particular the cols of P are the i,j,k coords of the new basis vectors. Then P converts a column of *new* coords of a point back to a column of *old* coords:

$$P \begin{bmatrix} \text{new} \\ \text{coords} \\ \text{of} \\ \text{a} \\ \text{vector} \end{bmatrix} = \begin{bmatrix} \text{old} \\ \text{coords} \\ \text{of} \\ \text{the} \\ \text{vector} \end{bmatrix}$$

And P^{-1} converts *old* coords to *new* (it will be shown in Section 3.1 that P is always invertible):

$$P^{-1} \begin{bmatrix} \text{old} \\ \text{coords} \\ \text{of} \\ \text{a} \\ \text{vector} \end{bmatrix} = \begin{bmatrix} \text{new} \\ \text{coords} \\ \text{of} \\ \text{the} \\ \text{vector} \end{bmatrix}$$

I'll call P the *basis-changing matrix*.

example 1

Let q = (3,5). Look at new basis u = (1,1), v = (4,3).
(a) One way to find the new coordinates of q w.r.t. basis u,v is to solve the vector equation q = au + bv. Do it instead by using the basis changing matrix.
(b) Find the connection in general between the old coordinates x,y of a vector and its new coordinates X,Y (express the new in terms of the old).

solution (a) Let $P = \begin{bmatrix} 1 & 4 \\ 1 & 3 \end{bmatrix}$. Then

$$P^{-1} = \frac{1}{|P|} \begin{bmatrix} 3 & -4 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ 1 & -1 \end{bmatrix}$$

$$P^{-1} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 11 \\ -2 \end{bmatrix}$$

So the new coordinates of q are 11,−2; i.e., p = 11u − 2v.
Here's a check: 11u − 2v = 11(1,1) − 2(4,3) = (3,5) which *is* q.

(b) $\begin{bmatrix} X \\ Y \end{bmatrix} = P^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3x + 4y \\ x - y \end{bmatrix}$

So

$$X = -3x + 4y$$
$$Y = x - y$$

new coordinates of the new basis vectors themselves

If u,v,w,q is a new basis for R^4 then

$$u = 1u + 0v + 0w + 0q$$
$$v = 0u + 1v + 0w + 0q$$
etc.

So the coords of u w.r.t. the basis u,v,w,p are 1,0,0,0; the coords of v w.r.t. the basis u,v,w,p are 0,1,0,0 etc. (It's silly to use the basis changing matrix in this easy case.)

example 2

Let

$$Z = \begin{bmatrix} 6 & 4 \\ 7 & 5 \end{bmatrix}$$

Suppose Z converts the coordinates of a vector w.r.t. basis i,j to the coords of the vector w.r.t. basis u,v.
(a) Let q = (1,10). Find the coords of q w.r.t. basis u,v.
(b) Find u and v (i.e., find their coordinates w.r.t. i,j).

solution

(a) $Z \begin{bmatrix} 1 \\ 10 \end{bmatrix} = \begin{bmatrix} 46 \\ 57 \end{bmatrix}$ so q = 46u + 47v.

(b) *method 1*

Z converts from old coords so Z^{-1} converts from new coords to old coords. So Z^{-1} is the basis changing matrix. The cols of Z^{-1} are the basis vectors u and v.

$$Z^{-1} = \frac{1}{|Z|} \begin{bmatrix} 5 & -4 \\ -2 & 5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 5 & -4 \\ -2 & 5 \end{bmatrix} = \begin{bmatrix} 5/2 & -2 \\ -1 & 3 \end{bmatrix}$$

So $u = (\frac{5}{2}, -1) = \frac{5}{2} i - j$ and $v = (-2, 3) = -2i + 3j$.

method 2 As in method 1, Z^{-1} converts the coords of a vector w.r.t. u,v to the coords of that vector w.r.t. i,j.

The coordinates of u w.r.t. u,v are 1,0. Convert them:

$$Z^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5/2 \\ -1 \end{bmatrix}$$

So the coords of u w.r.t. i,j are 5/2 and 1, i.e., $u = (\frac{5}{2}\ , 1) = \frac{5}{2}\ i + j$.

Similarly, the coords of v w.r.t. u,v are 0,1. Convert them:

$$Z^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

So the coords of v w.r.t. i,j are −2,3; i.e., v = (−2, 3) = −2i + 3j.

warning

In example 2 it is true by hypothesis that

$$Z\begin{bmatrix} \text{i-coord of vector q} \\ \text{j-coord of vector q} \end{bmatrix} = \begin{bmatrix} \text{u-coord of vector q} \\ \text{v-coord of vector q} \end{bmatrix}$$

But it is *not* true that

$$Z\begin{bmatrix} \text{vector i} \\ \text{vector j} \end{bmatrix} = Z\begin{bmatrix} \text{vector u} \\ \text{vector v} \end{bmatrix} \qquad \text{NO NO NO}$$

So it is *not* true in part (b) that u = 6i + 4j and v = 7i + 5j.

finding the new basis vectors when you know how old coordinates and new coordinates are related

Example 1 illustrated how to use the basis changing matrix to find the new coordinates of a vector.

Now suppose you know how the old coords x,y and the new coords X,Y are related and you want to find the basis vectors. In particular, suppose

(4) $$\begin{aligned} X &= 2x - 3y \\ Y &= y \end{aligned}$$

I'll find the new basis and draw a picture of the new coord system. Rewrite (4) as

$$\underbrace{\begin{bmatrix} X \\ Y \end{bmatrix}}_{\text{new coords}} = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\text{old coords}}$$

This means that

$$P^{-1} = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$$

Invert to get

$$P = \begin{bmatrix} 1/2 & 3/2 \\ 0 & 1 \end{bmatrix}$$

The new basis vectors u and v are the columns of P:

$$u = (\tfrac{1}{2},0)\,,\ v = (\tfrac{3}{2},1)$$

The new coord system is in Fig 1.
The X-axis lies on top of the x-axis but the X-scale is half the old scale .
Note that the fact that Y = y does *not* mean that the Y-axis is the same as the old Y-axis.
To check, let A have coords x = 4, y = 1 (Fig 1). By (4), the new coords of A are X = 5, Y = 1. Look at point A to see that its X and Y coords do look like 5 and 1.

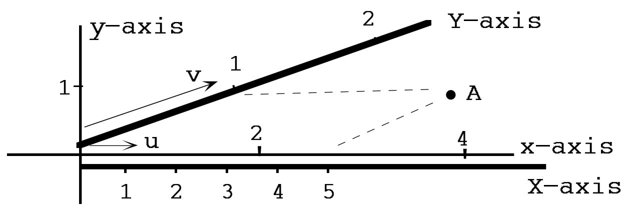


FIG 1

warning

When old coordinates x,y and new coordinates X,Y are related by

$$X = 2x - 3y$$
$$Y = y$$

it is *not* the case that the new basis vectors u,v and the old basis vectors i,j are related by

$$u = 2i - 3j \qquad \textit{wrong}$$
$$v = j \qquad \textit{wrong}$$

Instead it turned out that $u = \frac{1}{2}\,i$ and $v = \frac{3}{2}\,i + j$ and it took work to find this.

special case where only the scale changes

Suppose the new coord system in R^2 has the same axes as the old system but

(5) $$\text{X-scale} = 3 \times \text{x-scale}$$
$$\text{Y-scale} = \frac{1}{2} \times \text{y-scale}$$

(It's as if you went from using feet in the old coord system to using yards on the new X-axis and the half-foot on the new Y-axis.)
Fig 2 shows the old coord system and Fig 3 shows the new coord system.
The new basis vector u along the X-axis is 3 times as long as the old basis vector. And the new basis vector v on the Y-axis is half as long as the old basic vector. So

(6) $$u = 3i \text{ and } v = \frac{1}{2}\,j$$

The point A with coords x=1, y=0 has new coords X=1/3, Y=0.
The point B with coords x=0, y=1 has new coords X=0, Y=2.
In general, the connection between X,Y and x,y is

(7) $$X = \frac{1}{3}\,x,\; Y = 2y.$$

When only the scales change, halving a scale will halve the basis vector and double the coordinate; tripling a scale will triple the basis vector and one-third the coordinate.

The same idea works in R^n.

Here's an algebraic check that (6) and (7) go together.

With the new basis in (6), $P = \begin{bmatrix} 3 & 0 \\ 0 & 1/2 \end{bmatrix}$. Then $P^{-1} = \begin{bmatrix} 1/3 & 0 \\ 0 & 2 \end{bmatrix}$ and

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1/3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

So $X = \frac{1}{3} x$, $Y = 2y$.

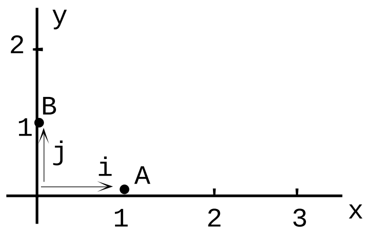


FIG 2

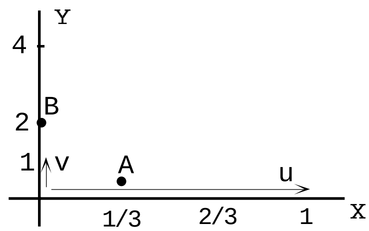


FIG 3

example 3

Suppose a point has coordinates X = 3, Y = 4 w.r.t. basis vectors $u = (1,3)$, $v = (1/2,1/2)$.

Find the coordinates of the point w.r.t. basis vectors u_{unit}, v_{unit}.

solution

$$u_{unit} = \frac{1}{\sqrt{10}} u$$

$$v_{unit} = \sqrt{2}\, v$$

Switching to the new basis vectors amounts to changing scales. In in a new X_1, Y_1 coord system with basis vectors u_{unit}, v_{unit}, we have $X_1 = \sqrt{10}\, X$, $Y_1 = \frac{1}{\sqrt{2}} Y$. So the new coords of the point are $X_1 = 3\sqrt{10}$, $Y_1 = 4/\sqrt{2}$.

example4 (changing from a new basis to a newer basis)

Let $u = (4,1,2)$, $v = (5,6,7)$, $w = (1,3,0)$ be a new basis for R^3.

Let $r = (1,0,0)$, $s = (0,2,5)$, $t = (1,1,1)$ be a newer basis for R^3.

Find a matrix that converts from coords w.r.t. u,v,w to to coords w.r.t. r,s,t.

solution First find a matrix that converts from u,v,w coordinates to i,j,k coordinates. By (3), the matrix that does this is

$$P = [u\ v\ w] = \begin{bmatrix} 4 & 5 & 1 \\ 1 & 6 & 3 \\ 2 & 7 & 0 \end{bmatrix}$$

Then find a matrix that converts from i,j,k to r,s,t. The matrix that converts from r,s,t coordinates to i,j,k coordinates is

$$Q = [r\ s\ t] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 5 & 1 \end{bmatrix}$$

Then Q^{-1} converts from i,j,k to r,s,t.

So far we have

$$P\begin{bmatrix}\text{coords}\\ \text{w.r.t}\\ \text{u,v,w}\end{bmatrix} = \begin{bmatrix}\text{coords}\\ \text{w.r.t.}\\ \text{i,j,k}\end{bmatrix} \quad\text{and}\quad Q^{-1}\begin{bmatrix}\text{coords}\\ \text{w.r.t.}\\ \text{i,j,k}\end{bmatrix} = \begin{bmatrix}\text{coords}\\ \text{w.r.t.}\\ \text{r,s,t}\end{bmatrix}$$

Put this together to get

$$Q^{-1}\,P\begin{bmatrix}\text{coords}\\ \text{w.r.t}\\ \text{u,v,w}\end{bmatrix} = \begin{bmatrix}\text{coords}\\ \text{w.r.t.}\\ \text{r,s,t}\end{bmatrix}$$

So $Q^{-1}P$ converts from u,v,w to r,s,t.

For example, suppose

(8) $$q = u - v + 2w.$$

To find the coords of q w.r.t. basis r,s,t, first I used a calculator to get

$$Q^{-1}P = \begin{bmatrix} 11/3 & -1/3 & -4 \\ 1/3 & 1/3 & -1 \\ 1/3 & 16/3 & 5 \end{bmatrix}$$

Then

$$Q^{-1}P\begin{bmatrix}1\\-1\\2\end{bmatrix} = \begin{bmatrix}-4\\-2\\5\end{bmatrix}$$

so the coordinates of q w.r.t. basis r,s,t are −4,−2,5, i.e.,

(9) $$q = -4r - 2s + 5t.$$

Check: $u - v + 2w = (4,1,2) - (5,6,7) + 2(1,3,0) = (1,1,5)$
$-4r - 2s + 5t = -4(1,0,0) - 2(0,2,5) + 5(1,1,1) = (1,1,5)$.

So the two representations of q in (8) and (9) agree.

footnote (more general basis changing rule)
The basis changing rule in (3) works with any two bases as follows, not just if one of the bases happens to be i,j,k.

Suppose you have two bases for R^n.
Call one of them old and one of them new (doesn't matter which is called the old and which is called the new as long as you stay faithful to your choice).

Let the first column of P be the old coordinates of the first new basis vector.
Let the second column of P be the old coordinates of the second new basis vector.
etc.

Then

$$P\begin{bmatrix}\text{new}\\ \text{coords}\\ \text{of a}\\ \text{vector}\end{bmatrix} = \begin{bmatrix}\text{old}\\ \text{coords}\\ \text{of that}\\ \text{vector}\end{bmatrix} \qquad \text{(P goes "backwards")}$$

And P^{-1} can be used to convert old coordinates to new coordinates.
I didn't use this in this last example when I converted from u,v,w

coordinates to r,s,t coords because in that example I was given the new basis vectors r,s,t in terms of i,j,k, not in terms of u,v,w. So I had to go roundabout.

coords of a vector w.r.t. an *orthogonal* basis

Suppose $\vec{u_1}, \ldots, \vec{u_n}$ is an orthogonal basis for R^n.

Let $\vec{x}$ be in R^n.

You can use the basis changing matrix to find the new coords of $\vec{x}$ but for an orthog basis there's an easier way. Here's the formula for doing it.

$$\boxed{\vec{x} = \frac{u_1 \cdot x}{u_1 \cdot u_1}\,\vec{u_1} + \ldots + \frac{u_n \cdot x}{u_n \cdot u_n}\,\vec{u_n}} \tag{10}$$

In the even more special case that the basis is ortho*normal*, the formula in (10) becomes

$$\boxed{\vec{x} = (\vec{u_1} \cdot \vec{x})\,\vec{u_1} + \ldots + (\vec{u_n} \cdot \vec{x})\,\vec{u_n}} \tag{11}$$

Here's the way to picture (10) in R^2.

If $\vec{u}, \vec{v}$ is an orthogonal basis for R^2 and $\vec{x}$ is in R^2 then (10) becomes

$$x = \frac{u \cdot x}{u \cdot u}\,u + \frac{v \cdot x}{v \cdot v}\,v$$

Remember that in R^2, the vectors $\frac{u \cdot x}{u \cdot u}\,u$ and $\frac{v \cdot x}{v \cdot v}\,v$ are the vector projections (Fig 4) of x onto u and v respectively (see Section 2.1). Their sum is x provided that u and v are perp.

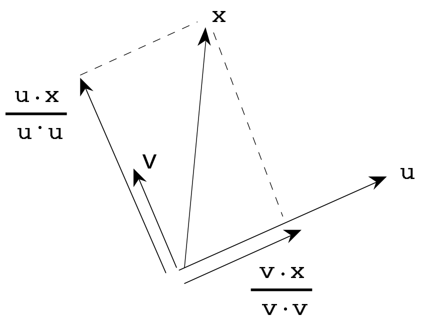


FIG 4

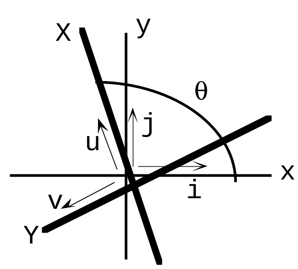


FIG 5

proof of (10)

Let $\vec{u_1},\dots,\vec{u_n}$ be an orthogonal basis for R^n. Let

(12) $$\vec{x} = x_1\vec{u_1} + x_2\vec{u_2} + x_3\vec{u_3} + \dots + x_n\vec{u_n}$$

I want to find the coefficients $x_1,\dots,x_n$.

Dot on both sides of (12) with say $\vec{u_2}$ to get

$$\vec{u_2}\cdot\vec{x} = \vec{u_2}\cdot(x_1\vec{u_1} + x_2\vec{u_2} + x_3\vec{u_3} + \dots + x_n\vec{u_n})$$

$$= x_1(\vec{u_2}\cdot\vec{u_1}) + x_2(\vec{u_2}\cdot\vec{u_2}) + x_3(\vec{u_2}\cdot\vec{u_3}) + \dots + x_n(\vec{u_2}\cdot\vec{u_n}) \quad \text{dot rules}$$

$$= x_2(\vec{u_2}\cdot\vec{u_2})$$ the other dot products are 0 since $u_1,\dots,u_n$ are orthogonal

So

$$x_2 = \frac{\vec{u_2}\cdot\vec{x}}{\vec{u_2}\cdot\vec{u_2}}$$

Similarly for all the other coeffs. This gives the formula in (10). QED

proof of (11)

If $\vec{u_1},\dots,\vec{u_n}$ are ortho*normal* then for each basis vector, $\vec{u_i}\cdot\vec{u_i} = \|\vec{u_i}\|^2 = 1$ so each denominator in (10) is 1.

rotating the standard coord system in R^2

If you get a new coord system in R^2 by rotating the old one counterclockwise by angle θ then the new orthonormal basis vectors (Fig 5) are

(13) $$u = (\cos\theta, \sin\theta)$$
$$v = (\cos[\theta+90^\circ], \sin[\theta+90^\circ]) = (-\sin\theta, \cos\theta)$$

Remember from §2.1 that $(r\cos\theta, r\sin\theta)$ is an arrow with length r, inclined at angle θ. I'm using $r = 1$ to get unit length vectors.

example 5

Rotate the axes in R^2 counterclockwise by 45° and find the new coords of the point $x = (4,2)$.

solution The new basis vectors are

$$u = (\cos 45^\circ, \sin 45^\circ) = (\tfrac{1}{2}\sqrt{2}, \tfrac{1}{2}\sqrt{2})$$
$$v = (-\tfrac{1}{2}\sqrt{2}, \tfrac{1}{2}\sqrt{2})$$

Since the new basis is orthonormal you can use the formula in (11):

$$x = (u\cdot x)u + (v\cdot x)v = 3\sqrt{2}\,u - \sqrt{2}\,v.$$

The new coords of x are $3\sqrt{2}, -\sqrt{2}$.

warning

In example 5, don't write $x = (3\sqrt{2}, -\sqrt{2})$ since that notation means coords w.r.t. the *original* coord system.

summary of some important ideas (you should know the answers to these questions)

question 1 Suppose r,s,t is a basis for R^3. What does it mean to say that matrix B converts from r,s,t coordinates to i,j,k coordinates.

answer It means that $B \begin{bmatrix} \text{coords} \\ \text{of a} \\ \text{vector} \\ \text{w.r.t.} \\ \text{r,s,t} \end{bmatrix} = \begin{bmatrix} \text{coords} \\ \text{of the} \\ \text{vector} \\ \text{w.r.t.} \\ \text{i,j,k} \end{bmatrix}$.

question 2 What does it mean to say that matrix B converts from r,s,t coordinates in R^3 to u,v,w coordinates.

answer It means that $B \begin{bmatrix} \text{coords} \\ \text{of a} \\ \text{vector} \\ \text{w.r.t.} \\ \text{r,s,t} \end{bmatrix} = \begin{bmatrix} \text{coords} \\ \text{of the} \\ \text{vector} \\ \text{w.r.t.} \\ \text{u,v,w} \end{bmatrix}$

PROBLEMS FOR SECTION 2.4

1. Let

$$u = 2i + j$$
$$v = 2i + 3j$$

be a new basis for R^2.

(a) Use matrices to find the components of (4,8) w.r.t. u and v.

(b) Check that your answer in (a) is right.

(c) Use matrices to find the components of i and j w.r.t. the new basis.

(d) Find the connection in general between old coords x,y and new coords X,Y (i.e., express x,y in terms of X,Y and vice versa.)

2. Let $\vec{x} = 4i + j - 2k$ and $w = 2i + 3j + 4k$

Find the coords of x w.r.t. the basis i,j,w

(a) by solving a system of equations

(b) and then again using the basis changing matrix

3. Suppose matrix B converts from i,j,k coordinates of a vector in R^3 to coordinates w.r.t. new basis u,v,w. What can you conclude if $B \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$.

4. Let

$$B = \begin{bmatrix} 1 & 2 & 1 \\ -1 & -1 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$

Suppose u,v,w is a basis for R^3 and B converts from i,j,k coords to u,v,w coords. Explain how to find u,v,w.

5. Here's how the old coords x,y (coords w.r.t. i,j) and new coords X,Y (w.r.t. a new basis u,v) are related. In each case, find the basis vectors u,v and sketch the new coord system.

(a) $x = 2X - Y$, $y = X + Y$

(b) $X = x + y$, $Y = 2x - y$

6. Define a new X,Y coord system by letting $X = 4x$, $Y = \frac{1}{2}y$. Find the new basis vectors u,v (a) by inspection (b) using fancy basis changing formulas

7. Look at point A = (1,3) in a standard coord system.
If the scale on each axis is halved (so that if the original scale is the foot then the new scale is the half-foot) find the new coords of A and the new basis vectors.

8. Let X and Y be the coordinates of a point w.r.t. a new basis u,v
(a) If X = 3x, Y = 2y, find u and v.
(b) If u = 4i and v = $\frac{1}{3}$ j find the equation of the circle $x^2 + y^2 = 5$ in the new coord system.

9. Let

u = 4i + 5j
v = 2i + 3j
p = 6i + 7j
q = 8i − j

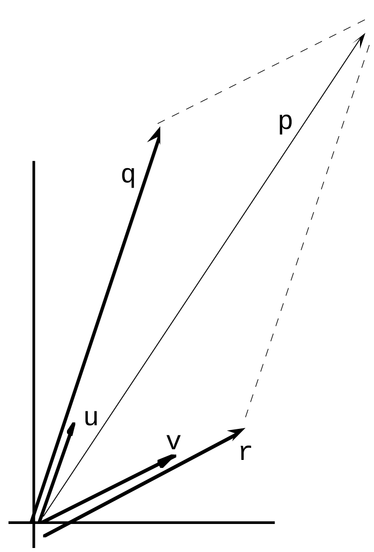
Find the matrix which converts
(a) from i,j coords to u,v coords
(b) from p,q coords to i,j coords
(c) from p,q coords to u,v coords
(d) from u,v coords to p,q coords

10. What does it mean to say that $\vec{p}$ has coordinates 2 and 3 w.r.t. the basis u,v.

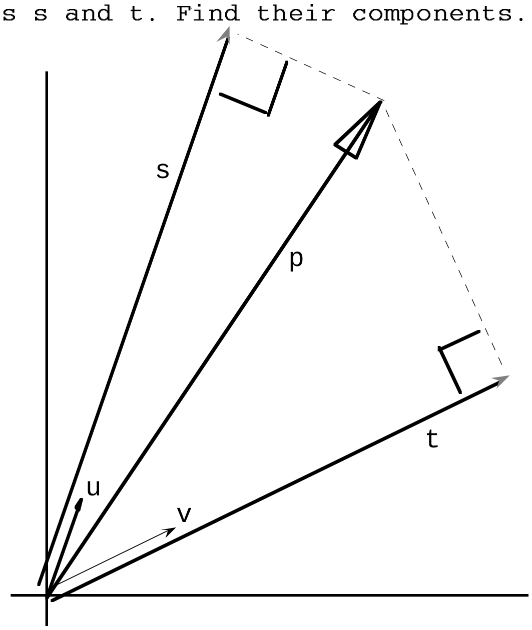
11. Let u = (3,2), v = (5,7).
Suppose the coords of p w.r.t. basis u,v are X = 2, Y = −1.
Find the old coords of p (w.r.t. basis i,j)
(a) using the basis changing matrix
(b) in some sensible fashion without resorting to a basis changing matrix

12. Let u = (−1,0), v = (0,2).
Let p = (2,6).
You should be able to find the new coords of p w.r.t. basis u,v by inspection.
But use overkill and find them
(a) with the basis changing matrix
(b) with a formula

13. Let u = (1,3), v = (4,2), p = (10,15).
(a) The diagram shows vectors q and r. (all arrows in the diagram are intended to start at the origin). Find their components (their usual components w.r.t. i,j).
(b) The second diagram shows vectors s and t. Find their components.



Problem 13(a)



Problem 13(b)

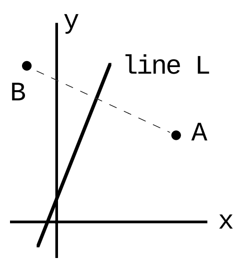
14. Find the new coords of point $p = (6,7)$ if the axes are rotated ccl by 29° (leave your answer in terms of cosines and sines).

15. If u,v,w,r are orthonormal vectors in R^4 and p is in R^4 then what good is

$$\sqrt{(u\cdot p)^2 + (v\cdot p)^2 + (w\cdot p)^2 + (r\cdot p)^2}$$

16. The problem is to find a formula for reflecting a point across the line L with equation $y = 3x$. In other words, if $A = (a,b)$, I want a formula for the coordinates of its reflection, B (see the diagram).
One way to do it is to switch to a new orthog coord system in which L is one of the axes, say the X-axis, solve the problem in the new coord system (it's easy to reflect in a coordinate axis) and then switch your answer back to the old system.
In particular, go through the following steps.

step 1 Find a pair of basis vectors for a new orthogonal system in which L is the X-axis.
step 2 Find the coords of A in the new cord system.
step 3 Finding the coords of B in the new system.
step 4 Convert your answer from step 3 to get the coords of B in the old coord system.



Problem 16

17. A new basis is $u = (2,1,3)$, $v = (1,1,1)$, $w = (0,0,1)$.
Another new basis is $p = (\pi,2,\sqrt{3})$, $q = (0,1,2)$, $r = (5,5,5)$.
Find a matrix that converts u,v,w coords to p,q,r coords.
But don't bother doing the actual computing. For instance an answer like $ABC^{-1}D^T$ is OK as long as you say what A,B,C,D are.

18. Let $u = (1,1,0)$, $v = (-1,1,0)$, $w = (0,0,2)$.
(a) Show that u,v,w is an orthogonal basis for R^3.
(b) Let $x = (\pi, 3, 6)$. Take advantage of (a) to find the coords of x w.r.t. u,v,w .

SECTION 2.5 SUBSPACES

the span of a bunch of vectors

Start with say u,v,w,p in R^6.
The *span* of u,v,w,p is the set of all combinations of u,v,w,p.
In other words, the span of u,v,w,p is the set of all vectors of the form au + bv + cw + dp. It contains u, v, w, p, 2u, 6u+3v+4w−6p, v−w, etc.

If the vectors are dependent then some of them can be thrown away without changing the span. In particular, if you pare down to a maximal number of independent spanning vectors, you don't change the span.

For example, if

$$u = (1,0,0,0,0,0),\ v = (2,3,0,0,0,0),\ w = (4,5,0,0,0,0),\ p = (7,0,0,0,0,0)$$

then u,v,w,p are dependent and every trio of them is also dependent but u and v are independent. The span of u,v,w,p is the same as the span of just u and v. It's also the same as the span of just w and p. And the same as the span of of just u and w. The span consists of all 6-tuples whose last four coordinates are 0.

The span of u and p is smaller than the original span. It is the same as the span of just u and consists of all vectors in R^6 whose last *five* coords are 0.

subspaces of R^n

Start with vectors $\vec{u}_1, \ldots, \vec{u}_k$ (not necessarily independent) in R^n.

Their span is called a *subspace* of R^n.

If $\vec{u}_1, \ldots, \vec{u}_k$ are *independent* then every vector in the subspace can not only be written as a combination of $\vec{u}_1, \ldots, \vec{u}_k$ but (by the unique representation rule in Section 2.2) can be written in only one way as a combination of $\vec{u}_1, \ldots, \vec{u}_k$. So we call $\vec{u}_1, \ldots, \vec{u}_k$ a basis for the subspace.

If $\vec{u}_1, \ldots, \vec{u}_k$ are *dependent* then you can get a basis for the subspace by paring them down to a maximal number of *independent* ones.

The dimension of a subspace is the number of vectors in a basis for the subspace, i.e., the number of independent spanning vectors.

As an extreme case, the subset of R^n containing only the vector $\vec{0}$ is a special case; it's considered to be a 0-dim subspace with no basis.

The other extreme case is R^n itself which is an n-dim subspace of R^n.

All other subspaces of R^n have dimension between 0 and n.

Many ideas about R^n also hold for subspaces of R^n:
(a) Once you know that the dimension of a subspace is say k then any set of k ind vectors in the subspace is a basis for the subspace.
(b) A set of more than k vectors in a k-dim subspace must be dependent.
(c) If u,v,w,p is an orthogonal basis for a subspace, and x is in the subspace, then

$$x = \frac{u\cdot x}{u\cdot u}\,u + \frac{v\cdot x}{v\cdot v}\,v + \frac{w\cdot x}{w\cdot w}\,w + \frac{p\cdot x}{p\cdot p}\,p$$

and if the basis is orthonormal then

$$x = (u\cdot x)u + (v\cdot x)v + (w\cdot x)w + (p\cdot x)p$$

the zero vector

Every subspace contains $\vec{0}$ (because if the subspace is spanned by u,v,w,p then one of the vectors in the subspace is $0u + 0v + 0w + 0p = \vec{0}$).

example 1 (some subspaces of R^3)

(a) Look at the vectors u and v in Fig 1. The subspace they span is the set of all combinations of u and v. If you picture vectors as points, then the subspace is the plane through the origin and through points u and v, i.e., the plane through the origin determined by arrows u and v.

Since the vectors u and v are independent, they are a basis for the subspace and the subspace is 2-dim. You can make a coord system for the plane by using u to determine an X-axis and v to determine a Y-axis.

A point in the plane has 3 coords w.r.t. i,j,k in R^3 but if you use the u,v coord system, the point can be described with only 2 coords. In Fig 1, the point $-u + 2v$ is in the plane and its u,v coords are $-1,2$.

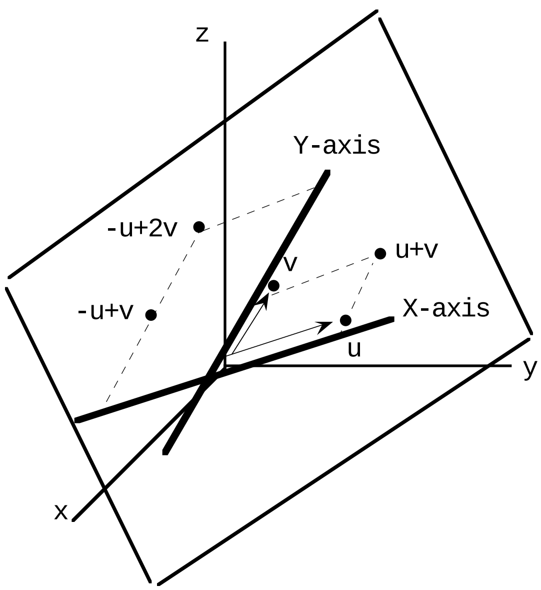
Any two independent vectors in the plane can also serve as a basis for the subspace.

(b) Look at the vector u in Fig 2. The subspace it spans is the set of all multiples of u. If you picture vectors as points, the subspace is the line through the origin in Fig 2 determined by the arrow u.

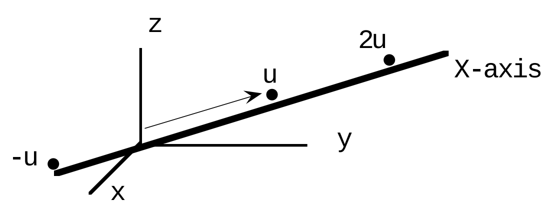
You can make a coord system for the line by using u to determine an X-axis.

A point on the line has 3 coords w.r.t. i,j,k but if you use the u coord system, the point can be described with only one coord.

The line is a 1-dim subspace of R^3, with a basis consisting of just the vector u.



2-dim subspace of R^3
FIG 1



1-dim subspace of R^3
FIG 2

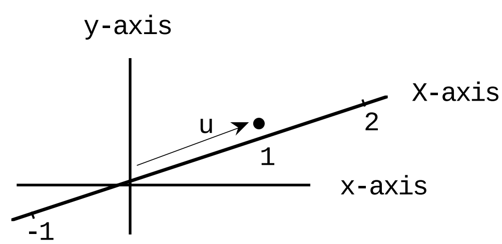


FIG 3

example 2 (a subspace of R^2)

Look at the vector u in Fig 3. The subspace it spans is the line through the origin determined by the arrow u.

You can make a coord system for the line by using u to determine an X-axis.

A point on the line has 2 coords w.r.t. i,j but if you use the u coord system, the point can be described with only one coord.

The line is a 1-dim subspace of R^2, with a basis consisting of just the vector u.

subspaces of R^2 in general

(a) There is one 0-dim subspace of R^2, the set containing just $\vec{0}$.

(b) The 1-dim subspaces of R^2 are the lines through the origin (if you take all combs of a single vector you get a line through the origin.)

(c) The only 2-dim subspace of R^2 is R^2 itself (if you take all combs of two ind vectors in R^2 you get R^2).

There are no subspaces of R^2 of dimension higher than 2.

subspaces of R^3 in general

(a) There is one 0-dim subspace of R^3, the set containing just $\vec{0}$.

(b) The 1-dim subspaces of R^3 are the lines through the origin (if you take all combs of a single vector you get a line through the origin.)

(c) The 2-dim subspaces of R^3 are the planes through the origin (if you take all combs of two ind vectors you get a plane through the origin).

(d) The only 3-dim subspace of R^3 is R^3 itself.

There are no subspaces of R^3 of dimension higher than 3.

example 3

Look at the subspace of R^5 spanned by

$$u = (1,1,2,2,0)$$
$$v = (0,0,0,0,1)$$

The vectors u and v are ind so they are a basis for the subspace and the subspace is 2-dim.

The vectors in the subspace must be of the form

$$au + bv = (a, a, 2a, 2a, b) \tag{1}$$

In other words, the subspace consists of points of the form (x_1,x_2,x_3,x_4,x_5) where

$$\begin{aligned} x_1 &= \text{anything} \\ x_2 &= x_1 \\ x_3 &= 2x_1 \\ x_4 &= 2x_1 \\ x_5 &= \text{anything} \end{aligned} \tag{2}$$

If

$$p = (4,5,6,7,8)$$

then p is not in the subspace because it doesn't have the form in (1). Another way to see that p is not in the subspace is to try to solve the vector equation

$$p = au + bv;$$

i.e., solve the system of equations (four equations in two unknowns)

$$\begin{aligned} 4 &= a \\ 5 &= a \\ 6 &= 2a \\ 7 &= 2a \\ 8 &= b. \end{aligned}$$

There is no sol for a and b so p is not in the subspace.

If

$$q = (5,5,10,10,9)$$

then q is in the subspace since it does have the form in (1). By inspection, q can be written as 5u + 9v. The coords of q w.r.t. the standard basis for R^5 are 5,5,10,10,9; its coords w.r.t. the basis u,v in the subspace are 5,9.

Any two ind vectors in the subspace are a basis. Another basis for the subspace is

$$r = (8,8,16,16,3), \ s = (2,2,4,4,7).$$

To find the coords of q w.r.t. the basis r,s (not obvious by inspection) solve the equation

$$q = ar + bs,$$

i.e., solve

$$\begin{aligned} 5 &= 8a + 2b \\ 5 &= 8a + 2b \\ 10 &= 16a + 4b \\ 10 &= 16a + 4b \\ 9 &= 3a + 7b \end{aligned}$$

The solution is $a = \frac{17}{50}$, $b = \frac{57}{50}$. So $q = \frac{17}{50} r + \frac{57}{50} s$.

example 4

Look at the subspace of R^5 spanned by

$$\begin{aligned} u &= (1,1,2,2,0) \\ v &= (0,0,0,0,1) \\ w &= (1,1,2,2,1) \quad \text{(note that } w = u + v\text{)} \end{aligned}$$

Since w = u + v, the 3 spanning vectors u,v,w are dep. But u and v are ind so the subspace is 2-dim and it's the same subspace as in example 3.

warning

1. A plane through the origin in R^3 is a 2-dimensional subspace of R^3. It resembles R^2 (it's intuitively a 2-dim world buried in a 3-dim world) but it is not the *same* as R^2.

Similarly, a 4-dim subspace of R^{10} resembles R^4 in many respects. In particular, every point in the subspace has four coordinates w.r.t. a basis for the subspace. But the subspace is not the *same* as R^4. The points in the subspace are 10-tuples; the points in R^4 are 4-tuples. So don't refer to a 4-dim subspace of R^{10} as if it *were* R^4.

2. Don't use the word "dimension" loosely.

You can say that a subspace has dimension 5 which specifically means that any basis for the subspace consists of 5 vectors.

You can say that a vector is 5-dimensional which means specifically that it is a member of R^5, i.e., is a 5-tuple.

But on an exam, *don't use the word "dimension" in any other context because it hasn't been defined in any other context*. If you do, you are probably talking intuition rather than mathematics.

closure

Saying that a set of vectors is *closed under addition* means that if u and v are in the set then so is u + v.

Saying that a set of vectors is *closed under scalar multiplication* means that if u is in the set and k is any scalar then ku is also in the set.

connection between subspaces and closure

A subspace is closed under addition and scalar multiplication.

To see why, look at the subspace spanned say by u,v,w,p. Here's a list of the vectors in the subspace:

u
$2u - 7v + 13w - \pi p$
$7v - 2w$
etc.

If you add any two of the vectors on this list you still have a combination of u,v,w,p so the sum is also on the list.
And if you take k times any vector on this list you get another combination of u,v,w,p so the product is also on the list.

Conversely, if a set of vectors is closed under addition and scalar multiplication then it is a subspace, meaning it is the span of some bunch of vectors (proof omitted).

how to show that a set of vectors is a subspace

method 1 Show that the set is composed of all combinations of a bunch of vectors; i.e., show that it is the span of some vectors.
method 2 Show that it's closed under addition and scalar mult.

how to try to show that a set of vectors is *not* a subspace

method 1 See if you can find specific vectors u and v so that u and v are in the subset but u + v is not in the subset. Then the set is not closed under addition so it can't be a subspace.
method 2 Try to find a specific vector u and a specific number k so that u is in the subset but ku is not in the subset. Then the set is not closed under scalar multiplication so it can't be a subspace.

method 3 If the set doesn't contain $\vec{0}$ then it is not a subspace. (But if the set does contain $\vec{0}$ then you have no conclusion.)

Note. It's possible for a set to be closed under addition but not under scalar mult (or vice versa) in which case it isn't a subspace.

example 5

Look at the set of vectors of the form (a,b,c,−a).
(a) Show that it's a subspace.
(b) Find a basis for the subspace and the dimension of the subspace.
(c) The vector x = (2,3,4,−2) is in the set. Find its coords w.r.t. your basis from part(b)

solution (a) Here's one method.
Look at the sum of two typical vectors in the set:

$$(a,b,c,-a) + (d,e,f,-d) = (a+d,\ b+e,\ c+f,\ -a-d)$$

The sum is in the set since its first and fourth coords are opposites.
And look at k times a typical vector in the set:

$$k(a,b,c,-a) = (ka,\ kb,\ kc,\ -ka)$$

The product is in the set since its first and fourth coords are opposites.
The set is closed under addition and scalar mult so it's a subspace.
(b) The set happens to consist of all combinations of the vectors

$$u = (1,\ 0,\ 0,\ -1)$$
$$v = (0,\ 1,\ 0,\ 0)$$
$$w = (0,\ 0,\ 1,\ 0)$$

If you can't see this by inspection then here's a way to do it mechanically: The set consists of all vectors (x_1,x_2,x_3,x_4) such that

$$\begin{aligned} x_1 &= a \\ x_2 &= b \\ x_3 &= c \\ x_4 &= -a \end{aligned}$$

Rewrite this as

$$\begin{aligned} x_1 &= 1a + 0b + 0c \\ x_2 &= 0a + 1b + 0c \\ x_3 &= 0a + 0b + 1c \\ x_4 &= -1a + 0b + 0c \end{aligned}$$

and then rewrite again:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = a \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}}_{u} + b \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{v} + c \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{w}$$

This shows that the set of points consists of all combinations of u,v,w.

Note. This shows that the set is a subspace so it serves as another method for doing part (a). Part (b) kills two birds with one stone.

The spanning vectors u,v,w are independent. I can see this by inspection; no one vector can be a combination of the others because of their first 3 components:

$$u = (\boxed{1},\boxed{0},\boxed{0},-1),\ v = (\boxed{0},\boxed{1},\boxed{0},0),\ w = (\boxed{0},\boxed{0},\boxed{1},0)$$

So u,v,w are a basis for the subspace. The subspace is 3-dim.

(c) x = 2u + 3v + 4w so its coords are 2,3,4 w.r.t. basis u,v,w.

warning

1. Don't say "(a,b,c,−a) is a subspace".
What you should say is "the set of points of the form (a,b,c,−a) is a subspace".
2. To show that a set of vectors is a subspace, you don't have to show that it contains $\vec{0}$. You have to either show that it is closed under addition and scalar mult or show that it has spanning vectors (in which case it automatically contains $\vec{0}$).

example 6

Let V be the set of points of the form (x_1,x_2,x_3,x_4,x_5) where $x_3 = x_1 x_2$. Show that it isn't a subspace.

first answer If u = (4,5,20,0,0) then u is in V but 2u is (8,10,40,0,0) which is not in V. So V is not closed under scalar mult. So it isn't a subspace.

second answer If u = (4,5,20,0,0) and v = (1,2,2,π,−2) then u and v are in V but u+v is (5,7,22,π,−2) which is not in V. So V is not closed under addition. So it isn't a subspace.

warning

In example 4, where V is the set of all vectors of the form $(x_2, x_2, x_1x_2, x_4, x_5)$ do *not* write $V = (x_2, x_2, x_1x_2, x_4, x_5)$.

rule of logic

To show that something is *always true* (e.g., to prove that a set is closed as in example 5a) you have to give a *general* argument.

To show that something is *not always true*, i.e., to *disprove* a generality (e.g., to disprove closure as in example 6) you should *not* give a general argument; instead you should give a *specific* counterexample.

mathmatical catechism (you should know the answers to these questions)

question 1 What does it mean to say that a set of vectors V is spanned by u,v,w.
answer It means that V consists of all vectors of the form au + bv + cw, i.e., V is the set of all combinations of u,v,w.

question 2 What is the span of u,v,w,p.
answer The set of all combinations of u,v,w,p, i.e., the set of all vectors of the form au + bv + cw + dp.

question 3 What is a subspace of R^n.
first answer A set of vectors in R^n which consists of all combinations of a bunch of vectors; i.e., the span of some bunch of vectors.
second answer A set of vectors in R^n that is closed under addition and scalar mult.

question 3 What does it mean to say that u,v,w is a basis for a subspace.
first answer It means that u,v,w are in the subspace and every vector in the subspace can be written in exactly one way as a combination of u,v,w.
second answer It means that u,v,w are independent and span the subspace.

question 4 What is the dimension of a subspace.
first answer The number of vectors in a basis for the subspace.
second answer The number of independent spanning vectors for the subspace.

question 6 What does it mean to say a set of vectors is closed under addition.
answer It means that the sum of any two vectors in the set is also in the set.

PROBLEMS FOR SECTION 2.5

1. Let

u = (1,2,3,4,5)
v = (1,1,0,0,0)
w = (3,4,3,4,5) (note that w = u + 2v)

For each of the following subspaces of R^5:
(i) Find a basis for the subspace and its dimension.
(ii) Decide if p = (0,−2,−6,−8,−10) and q = (4,5,6,7,8) are in the subspace and if they are, find their coords w.r.t. your basis.
(iii) Find umpteen more bases for the subspace.
(a) subspace spanned by u
(b) subspace spanned by u and v
(c) subspace spanned by u,v,w

2. Let $\vec{u} = 2\vec{i} + 6\vec{j}$. Give a geometric description of the subspace of R^2 spanned by

(a) $\vec{u}$ (b) $\vec{i}$ (c) $\vec{u},\vec{i}$ (d) $\vec{u}, \vec{i}, \vec{j}$

3. Let

$$\vec{p} = 2\vec{i} + 3\vec{j} + 4\vec{k}$$
$$\vec{q} = \vec{i} + \vec{j} + \vec{k}$$

Give a geometric description of the subspace of R^3 spanned by

(a) $\vec{i}$ (b) $\vec{i},\vec{p}$ (c) $\vec{i},\vec{p},\ \vec{i}+\vec{p}$ (d) $\vec{i},\vec{j},\vec{p}$ (e) $\vec{i},\vec{j},\vec{p},\vec{q}$

4. True or False. If a subspace is spanned by $\overrightarrow{u_1}$, $\overrightarrow{u_2}$, $\overrightarrow{u_3}$, $\overrightarrow{u_4}$, $\overrightarrow{u_5}$ then it is a 5-dim subspace.

5. Suppose u,v,w,p span a 2-dim subspace of R^7. Decide if possible if the following sets of vectors are ind or dep.

(a) u,v,w,p (b) u,v,w (c) u,v (d) u+v, u−v, u+w

6. Suppose $\vec{u}$, $\vec{v}$, $\vec{w}$ is a basis for a 3-dim subspace of R^{18}. Show that if $\vec{x}$ is orthogonal to u,v,w then x is orthog to every vector in the subspace.

7. Can a subspace contain exactly 7 vectors.

8. Are the following subspaces. If so, find a basis and the dimension.

(a) the set of points of the form (a,b,c,d,2)
(b) the set of points of the form (a,a^2,a^3,a^4,a^5)
(c) the set of points (x_1,x_2,x_3,x_4,x_5) where $x_3 = 2x_1$ and $x_5 = 3x_2$
(d) the set of points (x_1,x_2,x_3,x_4,x_5) where $x_3 = 2 + x_1$

9. Look at the set of points in R^5 satisfying the parametric equations

$$\begin{aligned} x_1 &= 2r + 4s - t \\ x_2 &= 8r + 9s + 2t \\ x_3 &= 8s + t \\ x_4 &= r - s \\ x_5 &= r + s + t \end{aligned}$$

(The parameters are r,s,t.)
(a) Find some points in the set just to show you understand what parametric equations are.
(b) Show that the set of points is a subspace of R^5.
(c) How would you find the dimension of the subspace.

10. Show that the set of vectors of the form (a,a,a,a,a) is a subspace. Find its dimension and a basis and find the coords of v = (3,3,3,3,3) w.r.t. your basis.

11. (a) What's the difference between R^2 and a 2-dim subspace of R^9.

(b) What is the span of u,v,p,q (what is its definition).
(c) What's the difference between the following two statements.
   (1) V is the subspace spanned by u,v,w
   (2) V is the subspace with basis u,v,w.
(d) What's the difference between a sub*set* of R^5 and a sub*space* of R^5.
(e) If u,v,w,p span a subspace does that make them a basis.
(f) What does it mean to say that a vector q can be written uniquely in terms of u,v,w,p.

12. (a) What do you think it would mean to say that a set of vectors is closed under subtraction.
(b) True or False (and defend your answer). A subspace is closed under subtraction.
(c) What do you think it would mean to say that a set of vectors is closed under normalizing.
(c) True or False (and defend your answer). A subspace is closed under normalizing.
(d) True or False ???????? A subspace is closed under dotting.

13. Example 5 showed that

$$u = (1, 0, 0, -1)$$
$$v = (0, 1, 0, 0)$$
$$w = (0, 0, 1, 0)$$

is a basis for the subspace of vectors of the form (a,b,c,−a).

(a) Check that the basis is orthogonal
(b) Let x = (5,6,7,−5). You should be able to write x in terms of u,v,w, by inspection but for practice try doing it with the formula for coords w.r.t. an orthogonal basis.

14. (a) Find a subset of R^2 that is closed under addition but not under scalar mult.
(b) Find a subset of R^2 that is closed under scalar mult but not under addition.
(c) Find a subset of R^5 that is closed under addition but not under scalar mult.
(d) Find a subset of R^5 that is closed under scalar mult but not under addition.

15. Example 5 shows that the set of points in R^4 of the form (a,b,c,−a) is a 3-dim subspace of R^4. What about the leftovers, i.e., the set of points *not* of the form (a,b,c,−a). Is it a subspace and if so, find a basis and the dimension.

16. In example 3, I found a basis for the subspace of all vectors of the form (a,b,c,−a). It turned out to be

$$u = (1, 0, 0, -1)$$
$$v = (0, 1, 0, 0)$$
$$w = (0, 0, 1, 0)$$

Now find a dozen more bases for the subspace.

17. Suppose u,v,w is a basis for a subspace of R^5.
Show that u, v + w, v − w is also a basis for the subspace.

18. Suppose u,v,w span a subspace but they aren't independent so you can't call them a basis for the subspace. What *do* you call them?

19. Look at the set of points on a circle centered at the origin in R^2. Is it a subspace.

20. Let A and B be fixed 10 × 7 matrices. Write vectors in R^7 as columns so they can be left multiplied by A and B.
Look at the set of vectors x in R^7 such that Ax = Bx, i.e., the set of vectors x that A and B "do the same thing to". Show that this set is a subspace.

21. Let

$$u = (1,2,3,4),\ v = (5,6,-1,2),\ w = (6,8,2,6),\ x = (4,4,-4,-2).$$

To decide if x is in the subspace of R^4 spanned by u,v,w, I used Mathematica to solve the system of equations au + bv + cw = x.

```
u = {1,2,3,4}; v = {5,6,-1,2}; w = {6,8,2,6}; x = {4,4,-4,-2};

Solve[x == a u + b v + c w ,{a,b,c}]

{{a -> -1 - c, b -> 1 - c}}
```

(a) How many equations and how many unknowns in the system.
(b) How many solutions does the system have and what conclusions follow about x and the subspace?

22. Do these say the same thing.
   (1) V contains only combinations of u,v,w.
   (2) V is spanned by u,v,w.
   (3) V contains all combinations of u,v,w.

23. Let $u_1 = (2,3,4,5,0,0,0)$
$u_2 = (1,\pi,6,2,0,0,0)$
$u_3 = (\sqrt{2},\sqrt{3},\sqrt{5},\sqrt{6},0,0,0)$
$u_4 = (1,1,1,1,0,0,0)$
$u_5 = (\pi,2,\pi,3,0,0,0)$

Decide, practically by inspection, if $u_1,\ldots, u_5$ are ind or dep.

REVIEW PROBLEMS FOR CHAPTER 2

1. If $\|\vec{w}\| = 3$, simplify $(\vec{v} - [\vec{w}\cdot\vec{v}]\vec{w})\cdot\vec{w}$.

2. Suppose the vectors

$$(a_1,a_2,a_3,a_4)$$
$$(b_1,b_2,b_3,b_4)$$
$$(c_1,c_2,c_3,c_4)$$

are independent. Are the following true or false. If false, find a counterexample.

(a) If you throw away the 4th components then the new vectors (a_1,a_2,a_3), (b_1,b_2,b_3), (c_1,c_2,c_3) are also ind.

(b) If you add arbitrary 5th components, the new vectors (a_1,a_2,a_3,a_4,a_5), (b_1,b_2,b_3,b_4,b_5), (c_1,c_2,c_3,c_4,c_5) are also ind.

3. (a) Write the generalization of the Pythagorean theorem in vector language so that it is a statement about vectors in R^n.

(b) Then use vector algebra to show that it is true in R^n.

4. Let u,v,w be any three vectors in R^{29}. Think subspace and explain why these four vectors must be dependent:

$$2u + 6v,$$
$$\pi v + \sqrt{5}\,w$$
$$7v + \frac{2}{3}w$$
$$8u - \pi v + 2w$$

5. If you had a computer that could do nothing but solve systems of equations, how would you use it to answer each of the following----what equation(s) would you ask it to solve and what would you do with the solution.

(a) Given vectors u,v,w,p, in R^{29}, are they independent.

(b) Given basis u,v,w,p for R^4 and vector x in R^4, find the coords of vector x w.r.t. the basis.

(c) Given u,v,w,y in R^4, is y in the subspace spanned by u,v,w.

(d) Given vectors q,r,s, u,v,w in R^4, do q,r,s span the same subspace as u,v,w.

(e) Let $A = \begin{bmatrix} 1 & 4 & 3 & -1 \\ 0 & 6 & 5 & 2 \\ 17 & 1/2 & 4 & 3 \\ 2 & 5 & 2 & 5 \end{bmatrix}$. Is A is invertible.

6. Suppose the coordinates X,Y w.r.t. a new basis u,v are related to the standard coords x,y w.r.t. basis i,j by the equations

$$X = 2x + 3y$$
$$Y = 5x - 6y$$

Pick one of the following.

(i) u = 2i + 3j, v = 5i − 6j

(ii) u = (2,5), v = (3,−6)

(iii) none of the above

7. True or False and defend your answer.

If u,v,w,p are independent vectors in R^{30} then 2u, 3v, 4w, 5p are ind also.

8. Suppose u is *not* in a particular subspace. Can 3u be in that subspace.

9. Let u = (1,2,3), v = (4,5,6), w = (5,6,5). I checked that they are ind.

$$\text{Let } B = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 5 & 6 \\ 3 & 6 & 5 \end{bmatrix}$$

$$\text{Then } B \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 13 \\ 13 \end{bmatrix}$$

Draw some conclusion from this about coordinates of some vector. In particular, what are the numbers 3,−1,2,9,13,13 good for.

10. Let u = (3,4,5,6), v = (9,10,11,12), p = (1,2,3,4).
Look at the vector equation p = au + bv. What system of equations is it.

11. Let u = 2i + j, v = i + 3j be a new basis for R^2.
Let x,y be the coords of a vector w.r.t. basis i,j.
Let X,Y be the coords of the vector w.r.t basis u,v.

Express X and Y in terms of x and y.

12. Look at the set of points on and inside the "unit sphere" in R^4, i.e., the points in in R^4 whose distance from the origin is ≤ 1 (distance in R^n was defined in §2.1 so this is a legal idea). Is it a subspace.

CHAPTER 3 THE VECTOR SPACE R^n CONTINUED

SECTION 3.1 ROW OPERATIONS AGAIN

the row space and col space of a matrix

The space spanned by the row vectors of a matrix (i.e., the set of all combinations of the row vectors) is called its *row space*.

The space spanned by the col vectors of a matrix is called its *column space*.

row operation rules

(1) Row ops do not change relations (or lack of relations) among the cols.

For example, if col 1 = 5 col 2 originally, then it is still true after row ops; if the first three cols are ind then they remain ind after row ops.

(2) Row ops can change the col space. But they can't change the dimension of the col space because, by (1), they can't change the maximal number of ind cols.

(3) Row ops do not change the row space of a matrix.

(4) Row ops can change the relations among the rows.

For example it's possible for the first three rows of matrix to be independent but become dependent after row ops are performed.

However, row ops cannot change the maximal number of ind rows (because this is the dimension of the unchanged row space).

semi-proof of (1)

Suppose

$$\text{old matrix} = \begin{bmatrix} a & d & 5a+7d \\ b & e & 5b+7e \\ c & f & 5c+7f \end{bmatrix}$$

so that col 3 = 5 col 1 + 7 col 2. Do the row op $R_2 = 2R_2$ (same idea works for other types of row ops). Then

$$\text{new matrix} = \begin{bmatrix} a & d & 5a+7d \\ 2b & 2e & 10a+14d \\ c & f & 5c+7f \end{bmatrix}$$

It's still true that col 3 = 5 col 1 + col 2.

proof of (2)

By (1), the two sets of cols have the same relations so they must have the same maximal number of ind cols.

But the two column spaces don't have to be identical. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The col space of A is the x–axis in R^3. If rows 1 and 3 are interchanged then

$$\text{new matrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$$

and the new col space is the z-axis in R^3. The two col spaces are different.

semi-proof of (3)

Suppose

$$\text{old matrix} = \begin{bmatrix} \text{row } u \\ \text{row } v \\ \text{row } w \end{bmatrix}$$

The row space consists of all vectors of the form au + bv + cw.
Do the row op $R_2 = R_2 + 2R_1$. Then

$$\text{new matrix} = \begin{bmatrix} u \\ 2u+v \\ w \end{bmatrix}$$

The new row space consists of all vectors of the form Au + B(2u+v) + Cw.
To see why the two row spaces are the same, look, say, at the vector

$$7u + 3(2u+v) - w$$

from the new row space. It can be written as

$$13u + 3v - w$$

so it's in the old row space too. And look at

$$8u + 3v - 9w$$

from the old row space. It can be written as

$$2u + 3(2u + v) - 9w$$

so it's in the new row space. This illustrates that whatever is in one row space is also in the other. So the two row spaces are the same.

proof of (4)

Look at

$$\text{old matrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In this matrix, row 2 is twice row 1 and rows 1,3 are ind.

Do the row op $R_2 = \frac{1}{2}R_2$. Then

$$\text{new matrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In the new matrix, row 2 is no longer twice row 1.

Or take the old matrix and do the row op $R_2 \leftrightarrow R_3$. Then

$$\text{new matrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 2 & 4 & 6 & 8 \end{bmatrix}$$

Now row 2 is no longer twice row 1 and rows 1,3 are no longer ind.
So row ops can change row relations.

But the maximal number of ind rows can't change (although their location can change) because that number is the dimension of the row space and, by (3), the row space doesn't change.

getting information about original columns and rows by looking at the echelon columns and rows

(5) Each echelon column without a pivot is a combination of the preceding echelon cols with pivots. By row op rule (1), the same relations hold for the corresponding columns in the original matrix.

(6) It can be shown that the echelon cols with pivots are a maximal set of ind echelon cols. So the corresponding original columns are a maximal set of ind original cols.

For example, suppose the echelon form of A is

$$\begin{bmatrix} 0 & \boxed{1} & 2 & 0 & 0 & 2 & \pi \\ 0 & 0 & 0 & \boxed{1} & 0 & 3 & 8 \\ 0 & 0 & 0 & 0 & \boxed{1} & 4 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ (the pivots are boxed)

Then

$$\text{echelon col } 7 = \pi \text{ echelon col } 2 + 8 \text{ echelon col } 4 + 9 \text{ echelon col } 5$$

so in the original matrix A,

$$\text{col } 7 = \pi \text{ col } 2 + 8 \text{ col } 4 + 9 \text{ col } 5$$

In the original A, cols 1,4,5 are a maximal ind set. There are other trios of cols in A that are ind (e.g., cols 3,4,5) but every foursome of cols of A is dep.

But it's a little trickier for rows.

In any echelon form, the nonzero echelon rows are ind and you can't get a bigger set of ind echelon rows. In other words, the nonzero echelon rows are a maximal set of ind echelon rows.

By the row op rules, this maximal number carries back to the original matrix.

In the echelon form of A given above, only the first 3 echelon rows are nonzero. So the maximal number of independent rows in the original A is 3. The 5 original rows are dep, every foursome is dep, some triple of original rows is independent but you don't know which triple(s). It is not necessarily true that the *first* three rows in A are ind.

But there is a special case where you can draw a firm conclusion about rows.

(7) If all the echelon rows are nonzero then the original rows are ind.

how to find bases for the row space and col space of a matrix

Remember that you can get a basis for the subspace spanned by u,v,w,p but cutting down to a maximal number of independent spanning vectors.

(8) The nonzero *echelon* rows are a basis for the row space of the *original* matrix.

In the special case that *all* the echelon rows are nonzero then in addition, the original rows (being ind) are also a basis for the row space of the original matrix.

That's because the nonzero echelon rows are a maximal set of ind echelon rows so they are a basis for the echelon row space which, by (3), is the same as the original row space.

(8a) In the special case that *all* the echelon rows are nonzero then not only are the echelon rows a basis for the row space of the original but the original rows are also a basis for the row space of the original matrix.

That's because in this case, by (7), the original rows are independent.

(9) The cols in the *original* matrix M which correspond to the echelon cols with pivots are a basis for the col space of the *original* matrix.

That's because the echelon cols with pivots are a maximal set of independent echelon cols. And so, by (1), the corresponding original cols are a maximal independent set of original cols.

example 1

Let

$$A = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 & u_5 \\ v_1 & v_2 & v_3 & v_4 & v_5 \\ w_1 & w_2 & w_3 & w_4 & w_5 \end{bmatrix}$$

Suppose the echelon form of A is

$$\begin{bmatrix} 1 & 2 & 0 & 4 & 6 \\ 0 & 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The row space of A is a 2-dim subspace of R^5 with basis

$(1,2,0,4,6)$, $(0,0,1,3,5)$ (the nonzero echelon rows)

The col space of A is a 2-dim subspace of R^3 with basis

(u_1,v_1,w_1), (u_3,v_3,w_3),

the cols in A corresponding to the echelon cols with pivots.

You also know that echelon cols 2 and 4 are ind, so cols 2 and 4 in A are ind so another basis for the col space of A is (u_2,v_2,w_2), (u_4,v_4,w_4).

And still another basis for the col space is (u_3,v_3,w_3), (u_4,v_4,w_4), etc.

warning

1. It's the nonzero *echelon* rows that are the basis for the row space of A, not necessarily the corresponding rows in A itself (except in the special case that *all* the echelon rows are nonzero).
2. It is *not* the *echelon* cols with pivots that are the basis for the col space of A; it's the corresponding cols in A itself.

how to find a basis for a subspace given spanning vectors

Suppose a subspace is spanned by vectors u,v,w,p. Here are two ways to get a basis for the subspace.

method 1 Find the "standard" basis.

Line up the spanning vectors as *rows* of a matrix and row op into echelon form. The nonzero echelon rows are a basis for the subspace, called the *standard* basis.

As examples will show, it's easy to decide what's in the subspace by inspection of the standard basis. Furthermore, the standard basis is unique. If two people start with different spanning vectors for the same subspace they will end up with the same standard basis.

method 2 Extract a basis from among the original spanning vectors.

Line up the spanning vectors as *columns* of a matrix and row op into echelon form. The original cols corresponding to the echelon cols with pivots are a basis for the subspace.

how to tell if q is in the subspace spanned by u,v,w

method 1 Line up u,v,w,q (with q last) as columns. Row operate into echelon form. It should be easy to tell if the 4th echelon column is a combination of the first three echelon columns. If so, then q is the same combination of u,v,w so it is in the subspace spanned by u,v,w. Otherwise, q is not in the subspace.

method 2 Line up u,v,w (without q) as rows. Row operate into echelon form. The nonzero echelon rows are the standard basis for the subspace spanned by u,v,w. The components of the standard basis are nice enough so that you can tell by inspection if q is a combination of them (see Example 2c for how this works). If so, then q is in the subspace spanned by u,v,w. Otherwise, q is not in the subspace.

example 2

Look at the subspace of R^4 spanned by

$$u = (0,1,2,4)$$
$$v = (3,-1,-1,0)$$
$$w = (2,-3,-2,-8)$$
$$p = (1,0,2,2).$$

Let

$$q = (1,2,3,4)$$
$$y = (5,0,-10,2)$$

(a) Find the standard basis for the subspace.
(b) Use part (a) to find a neat description of the points (x_1,x_2,x_3,x_4) of the subspace.
(c) Find a basis for the subspace from among the spanning vectors u,v,w,p.
(d) Decide if x and y are in the subspace.

solution (a) Line up u,v,w,p as rows and row op into echelon form:

$$\text{original} = \begin{bmatrix} 0 & 1 & 2 & 4 \\ 3 & -1 & -1 & 0 \\ 2 & -3 & -2 & -8 \\ 1 & 0 & 2 & 2 \end{bmatrix}, \quad \text{echelon form} = \begin{bmatrix} 1 & 0 & 0 & 6/5 \\ 0 & 1 & 0 & 16/5 \\ 0 & 0 & 1 & 2/5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A basis for the subspace is

$$r = (1,0,0,\ 6/5\ )$$
$$s = (0,1,0,\ 16/5)$$
$$t = (0,0,1,\ 2/5)$$

(b) The subspace consists of all vectors of the form

$$a\vec{r} + b\vec{s} + c\vec{t} = (a,b,c,\ \frac{6}{5}\ a + \frac{16}{5}\ b + \frac{2}{5}\ c).$$

So a point (x_1,x_2,x_3,x_4) is in the subspace iff $x_4 = \frac{6}{5}\ x_1 + \frac{16}{5}\ x_2 + \frac{2}{5}\ x_3$.

(c) Line up u,v,w,p as cols and row op into echelon form:

$$\text{original} = \begin{bmatrix} 0 & 3 & 2 & 1 \\ 1 & -1 & -3 & 0 \\ 2 & -1 & -2 & 2 \\ 4 & 0 & -8 & 2 \end{bmatrix}, \quad \text{echelon form} = \begin{bmatrix} 1 & 0 & 0 & 3/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In the echelon form, cols 1,2,3 have pivots. So cols 1,2,3 in the original matrix are a basis for the original col space. So u,v,w is a basis for the subspace spanned by u,v,w,p. (Echelon cols 1,2,4 are ind also, so u,v,p is another basis. In fact, in this instance, any three of the four vectors u,v,w,p happen to be a basis.)

(d) *method 1* Use the standard basis r,s,t from part (a). Let's test q.
The only way to get a combination of r,s,t to produce the first three components of q is to use r+2s+3t. But r+2s+3t doesn't have the right 4th component. So q is not in the subspace.
Or take advantage of part (b). The point (1,2,3,4) does not satisfy the equation $x_4 = \frac{6}{5} x_1 + \frac{16}{5} x_2 + \frac{2}{5} x_3$. So the point is not in the subspace,
Let's test y. The only way to get a combination of r,s,t to produce the first three components of y is to use 5r − 10t. And this time, the 4th component comes out right, namely it's 2. So y = 5r−10t; y is in the subspace. In other words, the point (5,0,−10,2) does satisfy the equation $x_4 = \frac{6}{5} x_1 + \frac{16}{5} x_2 + \frac{2}{5} x_3$. So it's in the subspace.

method 2 Line up u,v,w,p,q,y and row operate into echelon form. But you can leave out p because I already know from part (b) that u,v,w are a basis for the subspace spanned by u,v,w,p.

$$\text{original} \quad \begin{array}{ccccc} u & v & w & q & y \\ 0 & 3 & 2 & 1 & 5 \\ 1 & -1 & -3 & 2 & 0 \\ 2 & -1 & -2 & 3 & -10 \\ 4 & 0 & -8 & 4 & 2 \end{array}$$

$$\text{echelon form} \quad \begin{array}{ccccc} 1 & 0 & 0 & 0 & -13/2 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -7/2 \\ 0 & 0 & 0 & 1 & 0 \end{array}$$

In the echelon form, col 4 is not a comb of cols 1,2,3 so back in the original, q is not a comb of u,v,w. So q is not in the subspace.
In the echelon form, col 5 is a combination of cols 1,2,3. So back in the original, y is a comb of u,v,w. So y is in the subspace. In particular $y = -\frac{13}{2} u + 4v - \frac{7}{2} w$.

belated proof that 5 vectors in R^4 must be dependent (Section 2.2, (2))
Line up the 5 vectors as cols, and row op into echelon form. Here are three typical echelon forms.

$$\begin{array}{ccccc} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \qquad \begin{array}{ccccc} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 7 \\ 0 & 0 & 0 & 1 & 8 \end{array} \qquad \begin{array}{ccccc} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}$$

It isn't possible for all the cols to have pivots because there isn't enough room. There must be at least one col without a pivot. But a col without a pivot is always a combination of the preceding cols with pivots so the 5 echelon cols must be dep. So the 5 original cols are dep (by row op rule (3)). So the five vectors in R^4 are dep.

belated proof that any n independent vectors are a basis for R^n (Section 2.3, (2b))

Suppose u,v,w,p are independent vectors in R^4.

Let q be any vector in R^4. I'll show that q can be written as a combination of u,v,w,p.

Line up u,v,w,p,q as cols and row operate into echelon form:

$$[u \quad v \quad w \quad p \quad q] \quad \text{row ops to} \quad \begin{matrix} 1 & 0 & 0 & 0 & . \\ 0 & 1 & 0 & 0 & . \\ 0 & 0 & 1 & 0 & . \\ 0 & 0 & 0 & 1 & . \end{matrix}$$

The first four cols must row op to cols with pivots because u,v,w,p are ind. No matter what those row ops do to q, in the echelon form col 5 is a combination of the first four cols. So the same things holds for the original cols, i.e., q is a combination of u,v,w,p.

So u,v,w,p span R^4.

Furthermore, since u,v,w,p are ind, by the unique represention rule in Section 2.2, every vector in R^4 can be written in only one way as a combination of them.

So u,v,w,p is a basis for R^4.

invertible rule

Suppose A is a square matrix. The following are equivalent, i.e., either all are true or all are false.

(1) A is invertible (non-singular).
(2) $|A| \neq 0$.
(3) Echelon form of A is I.
(4) Rows of A are independent.
(5) Cols of A are ind.

proof that (5) belongs on the invertible list

cols of A are ind iff the echelon cols all have pivots
iff the echelon form of A is I

so item (5) is equivalent to (3) and belongs on the list.

proof that (4) belongs on the invertible list

Let A be $n \times n$.

The rows of A are ind iff max number of ind rows in A is n
iff max number of ind rows in echelon form is n (by row op rule (4))
iff the echelon rows are all nonzero
iff the echelon form is I

So item (4) is equivalent to (3) and belongs on the list.

belated proof that the basis changing matrix is invertible (Section 2.4, (3))

The cols of the basis changing matrix P are ind since they are a set of basis vectors. So by the invertible list, P must invert.

PROBLEMS FOR SECTION 3.1

1. Here's the echelon form of a matrix M.

$$\begin{matrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

This is not M. This is only the echelon form of M.

Answer whichever of the following are possible.

All the questions are about the *original* M.

(a) Are the cols of M ind or dep. If dep, find a dependency relation.
(b) Are the rows of M ind or dep. If dep, find a dependency relation.
(c) Find a basis for the row space of M.
(d) Find a basis for the col space of M.
(e) Let x = (2,3,4,5,6). Is x in (i) the row space of M (ii) the col space of M
(f) Let y = (5,−3,1,0,0). Is y a combination of (i) the rows of M (ii) the cols of M
(g) Are rows 1,2,3 of M ind or dep.
(h) Are rows 1,2,3,4 of M ind or dep.
(i) Are cols 1,2,3 of M ind or dep.
(j) Are cols 2,3,4 of M ind or dep.

2. Suppose M has the indicated echelon form. In each case, fill in the question marks.

The col space of M is a ?-dim subspace of $R^?$ with basis ?

The row space of M is a ?-dim subspace of $R^?$ with basis ?

(a) $\begin{matrix} 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}$ (b) $\begin{matrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{matrix}$ (c) $\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$

3. Let

$$M = \begin{bmatrix} 1 & 2 & 3 & 1 & 2 & 0 \\ 2 & 4 & 6 & 2 & -2 & -1 \\ 2 & 4 & 6 & 1 & 4 & 1 \end{bmatrix}$$

(a) Find a basis for the col space.
(b) Find a basis for the row space.
(c) Let x = (5,6,7). Is x in the col space.
(d) Let y = (0,0,0,1,12,1). Is y in the row space.

4. Let

$$\begin{aligned} u &= (2,2,-4,5) \\ v &= (0,1,-4,4) \\ w &= (6,4,-4,7) \end{aligned}$$

(a) Are u,v,w ind. If not, find a dependency relation.
(b) Find a basis for and the dimension of the subspace spanned by u,v,w by paring down the set of spanning vectors.
(c) Let x = (3,4,5,6). Is x in the space spanned by u,v,w.

5. Given vectors u_1, u_2, u_3, u_4 in R^4.

Let A be the echelon form of the matrix with *cols* u_1, u_2, u_3, u_4 and let B be the echelon form of the matrix with *rows* u_1, u_2, u_3, u_4:

$$A = \begin{matrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix} \qquad B = \begin{matrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{matrix}$$

(a) Are the vectors u_1, u_2, u_3, u_4 ind or dep. If dep, find a dependence relation
(b) Find a basis for the subspace spanned by u_1, u_2, u_3, u_4 by selecting from among the spanning set. In fact find several such bases.
(c) Find the standard basis for the subspace spanned by u_1, u_2, u_3, u_4.
(d) Is u_4 in the subspace spanned by u_1, u_2, u_3. If so, find its coords w.r.t. a basis for the subspace.

(e) Is u_3 in the subspace spanned by u_1 and u_2. If so, find its coords w.r.t. a basis for the subspace.
(f) Is u_1 in the subspace spanned by u_2 and u_3.
(g) Let p = (2,3,4,5). Is p in the subspace spanned by u_1, u_2, u_3, u_4.

6. (a) Describe the row space of the matrix $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ and find a basis for it.

(b) Describe the col space of the matrix $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \end{bmatrix}$ and find a basis for it.

7. Let

u = (1,2,3,4)
v = (0,0,0,1)
w = (1,0,0,0)
p = (2,2,3,3)
q = (2,0,0,3)
r = (1,0,0,1)
s = (1,0,0,2)

The problem is to decide if the subspace spanned by u,v,w is the same as the subspace spanned by p,q,r,s .
(a) Do it using ideas from this section.
(b) Describe how to do it using a computer that could only solve systems of equations.

8. Let M be 5 $\times$ 5. Let u = (1,4,2,9,3). Can you tell if u is in the row space of M if
(a) M is invertible (b) M is not invertible

9. Construct a matrix with ind rows and dep cols.

10. In the "proof that (5) belongs on the invertible list", the word "iff" appears several times. Is that a typo?

11. Here is the incomplete list from Section 1.3 of things to look for in the rows (or cols) in order to spot a zero determinant:
(i) a row of zeroes
(ii) two identical rows
(iii) one row a multiple of another row

What would make the list complete.

12. On an exam, a question asked for a basis for a given subspace. The exam solutions gave the answer

u = (4,5,1,2,0), v = (6,2,2,7,−1), w = (1,0,−2,1,−2)

A student gave the answer

p = (2,8,0,−3,1), q = (3,8,−2,−2,−1), r = (−8,−7,−7,−7,−3)

Imagine yourself with a computer that can do row operations and explain how would you decide if her answer is OK by lining up vectors as
(a) rows
(b) columns

13. Let u = (1,2,0,3,0), v = (1,2,1,7,1), w = (0,0,0,0,1).
(a) Find neat equations that a point $(x_1, x_2, x_3, x_4, x_5)$ must satisfy to be in the subspace.
(b) Is $(2,4,-1,2,\pi)$ a combination of u,v,w.

SECTION 3.2 PROJECTIONS

projecting a point into a subspace

Given a subspace and a vector $\vec{q}$.

Here are two ways to describe the projection $\vec{q}_{proj}$ of $\vec{q}$ onto the subspace.

(1) The projection q_{proj} is the point in the subspace such that the vector $q-q_{proj}$ is orthogonal to every vector in the subspace.

Fig 1 illustrates the idea, with a 2-dim subspace of R^3.

(2) The projection q_{proj} is the point in the subspace that is closest to q; i.e., the minimum value of the norm $\|q - \text{point in subspace}\|$ occurs when the point is q_{proj}.

It can be shown that (1) and (2) are equivalent.
And it can be shown that the projection is unique.
Note that if q is in the subspace to begin with then q_{proj} is just q itself.

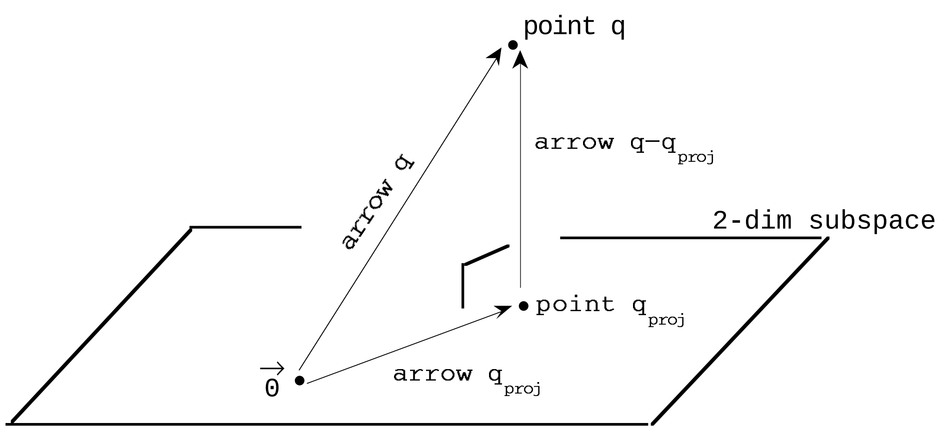


FIG 1

finding the projection given a basis (or just spanning vectors) for the subspace

Given a subspace of R^n with basis u,v,w and given q in R^n.

Let q_{proj} be the projection of q into the subspace.
Then $q_{proj} = au + bv + cw$.

To find a,b,c, make $q-q_{proj}$ orthog to every vector in the subspace. It's enough to make $q-q_{proj}$ orthog to the basis vectors u,v,w (a problem in Section 2.5 shows that if a vector is orthog to basis vectors, it's orthog to every vector in the subspace). So solve the equations

$$
\begin{aligned}
u\cdot(q-q_{proj}) &= 0 \\
v\cdot(q-q_{proj}) &= 0 \qquad (3)\\
w\cdot(q-q_{proj}) &= 0
\end{aligned}
$$

called the *normal* equations. The equations can be written in matrix form as follows. Let A be the matrix with cols u,v,w so that the rows of A\S(T) are u,v,w. Write q and q_{proj} as columns. Then (3) becomes

$$A^T(q - q_{proj}) = \vec{0}$$

$$A^T q - A^T q_{proj} = \vec{0}$$

$$A^T q = A^T q_{proj}$$

But

$$q_{proj} = au + bv + cw = a\begin{bmatrix} \text{col 1} \\ \text{of} \\ A \end{bmatrix} + b\begin{bmatrix} \text{col 2} \\ \text{of} \\ A \end{bmatrix} + c\begin{bmatrix} \text{col 3} \\ \text{of} \\ A \end{bmatrix} = A\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

(see observation (1) about matrix mult in Section 1.1). So finally (3) becomes

$$A^T q = A^T A \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

All of this works even if u,v,w span the subspace but are not a basis (i.e., this works if u,v,w span but are not independent). All in all:

Let u,v,w span a subspace of R^n. Let q be in R^n.
To find a,b,c so that $q_{proj} = au + bv + cw$, solve the normal equations

$$A^T A \begin{bmatrix} a \\ b \\ c \end{bmatrix} = A^T q \tag{4}$$

where A = [u v w] (i.e., the cols of A are u,v,w).

The projection of q into the subspace spanned by u,v,w is often referred to as the best approximation to q of the form au + bv + cw.

example 1

Let

u = (1,0,0,0)
v = (1,0,1,0)
w = (0,1,0,2)
q = (3,2,1,2)

I'll find the projection of q into the subspace spanned by u,v,w.

Let

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Then

$$A^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}, \qquad A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$A^T q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}$$

The normal equations in (4) become

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}$$

This is the system of equations

(5)
$$\begin{aligned} a + b &= 3 \\ a + 2b &= 4 \\ 5c &= 6 \end{aligned}$$

The solution is

$$c = 6/5, \quad a = 2, \quad b = 1$$

so

$$\begin{aligned} q_{proj} &= 2u + v + \tfrac{6}{5}\,w. \\ &= 2(1,0,0,0) + (1,0,1,0) + \tfrac{6}{5}(0,1,0,2) \\ &= (3, \tfrac{6}{5}, 1, \tfrac{12}{5}). \end{aligned}$$

So q_{proj} has coords 2, 1, 6/5 w.r.t. the basis u,v,w for the subspace and has coords 3, 6/5, 1, 12/5 w.r.t. basis i,j,k,ℓ for R^4.

warning

1. In example 1, after you solve the normal equations and get a=2,b,=1,c=6/5 you must know what to do with the solution. The projection is *not* (2,1,6/5). That makes no sense since q_{proj} is in R^4, not R^3. Rather, the projection is $2u + v + \frac{6}{5}w$. And it is (3, 6/5, 1, 2/5).

2. Don't use the word projection or the notation q_{proj} unless it is clear what subspace you are projecting into.

finding the projection given an orthogonal basis for the subspace

Suppose u,v,w is an orthogonal basis for a 3-dim subspace. Here's a formula for the projection of q into the subspace (same idea works for a k-dim subspace):

(6)
$$\boxed{q_{proj} = \frac{u\cdot q}{u\cdot u}\,u + \frac{v\cdot q}{v\cdot v}\,v + \frac{w\cdot q}{w\cdot w}\,w}$$

In the still more special case that u,v,w is an ortho*normal* basis, (6) turns into

(7)
$$\boxed{q_{proj} = (u\cdot q)u + (v\cdot q)v + (w\cdot q)w}$$

The formulas on the right side of (6) and (7) have dual use.

If u,v,w is an orthogonal basis for a subspace of R^n and q is in R^n then

$$\frac{u\cdot q}{u\cdot u}\,u + \frac{v\cdot q}{v\cdot v}\,v + \frac{w\cdot q}{w\cdot w}\,w = \begin{cases} q_{proj} & \text{if q is outside the subspace} \quad \text{((6) in this section)} \\ q & \text{if q is in the subspace} \quad \text{((10) in Section 2.4)} \end{cases}$$

And if u,v,w is an ortho*normal* basis for a subspace then

$$(u\cdot q)u + (v\cdot q)v + (w\cdot q)w = \begin{cases} q_{proj} & \text{if q is outside the subspace} \quad \text{((7) in this section)} \\ q & \text{if q is in the subspace} \quad \text{((11) in Section 2.4)} \end{cases}$$

Fig 3 shows how to picture the formula for projecting q onto a 2-dim subspace of R^3 with orthog basis u,v. Remember that in R^3, the vectors $\frac{u\cdot q}{u\cdot u}\,u$ and $\frac{v\cdot q}{v\cdot v}\,v$ are the vector projections of q onto u and v respectively (see Section 2.1). Their sum is q_{proj} provided that u and v are perp.

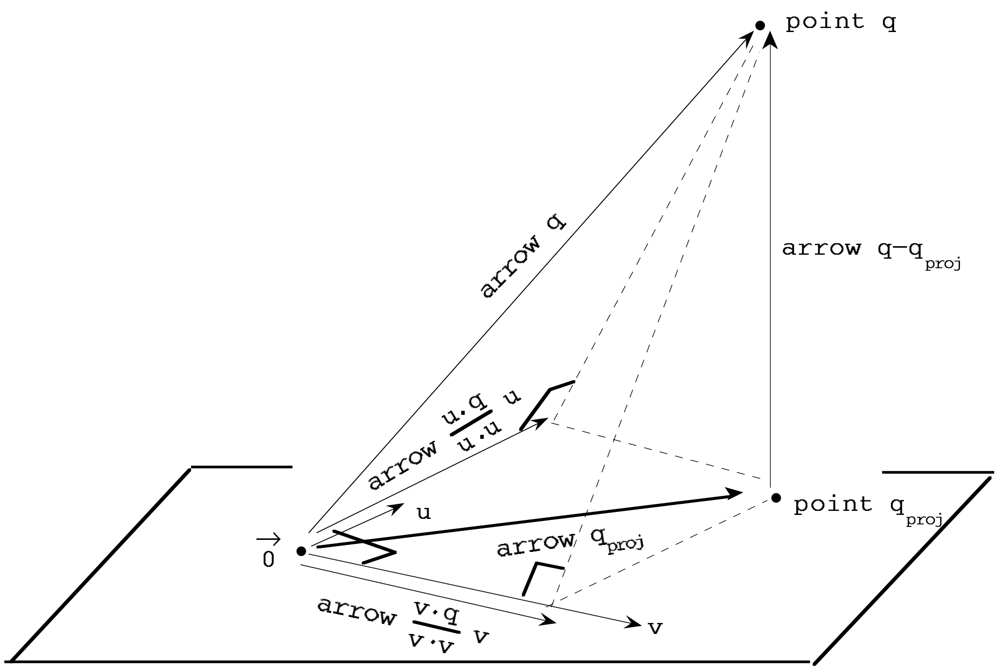


FIG 3

proof of (6)

Suppose u,v,w is an orthogonal basis for a 3-dim subspace of R^n.

The vector q_{proj} is in the subspace so by the rule for finding the coords of a vector w.r.t. an orthogonal basis (Section 2.4, (10)),

$$q_{proj} = \frac{u\cdot q_{proj}}{u\cdot u}\,u + \frac{v\cdot q_{proj}}{v\cdot v}\,v + \frac{\mathbf{w}\cdot q_{proj}}{\mathbf{w}\cdot\mathbf{w}}\,\mathbf{w} \tag{8}$$

By the description of a projection in (1), $q-q_{proj}$ is orthog to u so

$$u\cdot(q-q_{proj}) = 0$$
$$u\cdot q_{proj} = u\cdot q$$

So in (8), $\frac{u\cdot q_{proj}}{u\cdot u}\,u$ can be replaced by $\frac{u\cdot q}{u\cdot u}\,u$. If you stare at Fig 3, you can see the geometric interpretation of this: the projection of q onto u is the same as the projection of q_{proj} onto u.

Similarly,

$$v\cdot q_{proj} = v\cdot q$$
$$\mathbf{w}\cdot q_{proj} = \mathbf{w}\cdot q$$

and (8) turns into (6). QED

example 2

Find the projection of $q = (2,3,4)$ in the subspace of R^3 with orthonormal basis

$$u = \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

$$v = \left(-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$$

solution

$$q_{proj} = (u \cdot q)u + (v \cdot q)v = \frac{5}{\sqrt{3}}\,u - \frac{7}{\sqrt{6}}\,v = (2,-2,19).$$

So q_{proj} has coords $\frac{5}{\sqrt{3}}, -\frac{7}{\sqrt{6}}$ w.r.t. the basis u,v for the subspace and has standard coords 2,−2,19 w.r.t. basis i,j,k in R^3.

PROBLEMS FOR SECTION 3.2

1. Let $u = (1,1,0)$, $v = (2,5,0)$, $q = (4,3,9)$.
Find the projection of q onto the subspace spanned by u, v
 (a) by inspection using geometry
 (b) with the normal equations

2. Let $u = (2,-1,1,0)$, $v = (0,2,0,1)$, $x = (1,2,1,1)$.
 (a) Find the projection of x into the subspace spanned by u and v.
 (b) Is x_{proj} supposed to be perp to something? Check it out.

3. Let $u = (1,1,0)$, $v = (1,0,1)$, $y = (1,-1,1)$.
 Find the best approximation to y of the form au + bv.

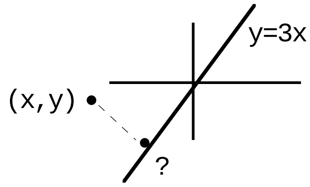
4. Let $u = (2,3,4)$, $v = (4,0,-2)$, $q = (1,2,3)$.
Find the projection of q into the subspace of R^3 with basis u,v (do it the easy way).
Find its coords w.r.t. basis u,v and find its standard coords (w.r.t. basis i,j,k).

5. Find the projection of $u = (2,3,6,7)$ onto the 1-dim subspace spanned by $v = (1,5,-2,1)$.

6. Let $p = (\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$, $q = (-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$, $u = (1,2,3)$.
Find the projection of u onto the subspace spanned by p,q. In particular, find its coords w.r.t. basis p,q and its coords w.r.t. the standard basis for R^3.
Do it three times:
(a) the easy way
(b) with the normal equations
(c) using max/min ideas from calculus

7. (a) Use the material in this section to find the projection of point (x,y) onto the line y = 3x (see the diagram).

(b) Why doesn't the material in this section apply to finding the projection of point (x,y) onto the line y = 3x + 7.



Problem 7(a)

8. Let u,v,w be orthonormal vectors in R^{17}. Let x be in R^{17}. Let

$$k = \sqrt{(u\cdot x)^2 + (v\cdot x)^2 + (w\cdot x)^2}$$

Decide what k represents and then use geometric intuition to predict which is larger, $\|x\|$ or k.

9. (a) Of all vectors of the form (a,2a,b,3b) find the one closest to (2,5,3,4).
(b) What does it mean to say that your answer to (a) is "closest" to (2,5,3,4).

10. Suppose $\vec{u_1}$, $\vec{u_2}$, $\vec{u_3}$, $\vec{u_4}$ are orthogonal vectors in R^4 and $\vec{x}$ is in R^4.
What do the following represent.

(a) $$\frac{u_1\cdot x}{u_1\cdot u_1}u_1 + \frac{u_2\cdot x}{u_2\cdot u_2}u_2 + \frac{u_3\cdot x}{u_3\cdot u_3}u_3 + \frac{u_4\cdot x}{u_4\cdot u_4}u_4$$

(b) $$\frac{u_2\cdot x}{u_2\cdot u_2}u_2 + \frac{u_3\cdot x}{u_3\cdot u_3}u_3 + \frac{u_4\cdot x}{u_4\cdot u_4}u_4$$ (same as (a) but missing the first term)

11. Suppose $\vec{u_1}$, $\vec{u_2}$, $\vec{u_3}$, $\vec{u_4}$ are orthogonal vectors in R^5 and $\vec{x}$ is in R^5.

(a) What does $\frac{u_1\cdot x}{u_1\cdot u_1}u_1 + \frac{u_2\cdot x}{u_2\cdot u_2}u_2 + \frac{u_3\cdot x}{u_3\cdot u_3}u_3 + \frac{u_4\cdot x}{u_4\cdot u_4}u_4$ represent.

(b) Fill in the blanks.
The vector in part (a) is the best approximation of ______ of the form ________.

(c) Fill in the blanks.
Of all vectors in _______ , the vector in part (a) is the one closest to _______.

12. Let V be the set of vectors in R^5 where the first component is the sum of the last two components. Let q = (1,2,3,4,5).
(a) Find the vector in V closest to q.
(b) Find several other vectors in V and check that the one you found in part (a) really is closer to q than those others.

13. Use ideas from this section to find the distance from the point (0, 2, −1) to the plane 2x + 2y + z = 0.

SECTION 3.3 LEAST SQUARES SOLUTION TO AN INCONSISTENT SYSTEM OF EQUATIONS

example 1

Look at the system

$$\begin{aligned} x + y &= 4 \\ 2x + 3y &= 1 \\ 2x + y &= 3 \end{aligned} \tag{1}$$

It is inconsistent (i.e., has no solution) since the last two equations need $y = -1$, $x = -5$ but these values don't satisfy the first equation. Here's a way to get the "best" non-solution.

first point of view The system of equations in (1) can be written in matrix form as

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}}_{\vec{b}} \tag{2}$$

It also can be written in vector form as

$$x \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{u} + y \underbrace{\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}}_{v} = \underbrace{\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}}_{b} \tag{3}$$

so the problem is to write b as xu + yv. This is just as impossible as the original problem but the next best thing is to find the combination of u and v that is *closest* to b, in other words, find scalars x and y so that

$$x\vec{u} + y\vec{v} = \vec{b}_{proj}$$

where b_{proj} means the projection of b onto the subspace spanned by u and v.

To get the normal equations for finding x and y, let

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix}$$

Then

$$A^T A = \begin{bmatrix} 9 & 9 \\ 9 & 11 \end{bmatrix}, \qquad A^T b = \begin{bmatrix} 12 \\ 10 \end{bmatrix}$$

The normal equations are

$$A^T A \vec{x} = A^T \vec{b}$$

Note that the normal equations happen to be the original equations in (2) but with an extra A^T on the left of each side. The normal equations are

$$\begin{bmatrix} 9 & 9 \\ 9 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 12 \\ 10 \end{bmatrix}$$

$$\begin{aligned} 9x + 9y &= 12 \\ 9x + 11y &= 10 \end{aligned}$$

The solution to the normal equations is

$$x = \frac{7}{3}, \quad y = -1$$

These values of x and y do not solve the original system of equations (*nothing* does). But $\frac{7}{3}u - v$ is the combination of u and v closest to b. So $x = \frac{7}{3}$, $y = -1$ is considered to be the best non-solution of the inconsistent system.

second point of view To solve (1) you want x+y to be 4, 2x+3y to be 1 and 2x+y to be 3. For any x,y,z, the "errors" are how far off x+y, 2x+3y and 2x+y actually are:

$$\begin{aligned} \text{error1} &= x + y - 4 \\ \text{error2} &= 2x + 3y - 1 \\ \text{error3} &= 2x + y - 3 \end{aligned} \tag{4}$$

Look at the sum of the squared errors (square the errors so that a positive error in the sum doesn't cancel a negative error):

$$\text{sum of square errors} = (x+y-4)^2 + (2x+3y-1)^2 + (2x+y-3)^2 \tag{5}$$

If x and y were a solution to (1) then the sum of the square errors would be zero. Since there is no solution, the next best thing is to choose x and y so as to minimize the sum of the square errors (this is a calculus max/min problem).

Here's how minimizing the sum of square errors leads back to the projection. In vector notation, with u,v,b as in (3), the three equations in (4) can be written as

$$\begin{bmatrix} \text{error1} \\ \text{error2} \\ \text{error3} \end{bmatrix} = x \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{\vec{u}} + y \underbrace{\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}}_{\vec{v}} - \underbrace{\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}}_{\vec{b}} \qquad = x\vec{u} + y\vec{v} - \vec{b}$$

and

$$\text{sum of square errors} = \|x\vec{u} + y\vec{v} - \vec{b}\|^2$$

Minimizing the sum of square errors is the same as minimizing $\|x\vec{u} + y\vec{v} - \vec{b}\|^2$ which can be done by minimizing $\|x\vec{u} + y\vec{v} - \vec{b}\|$ itself. This in turn is the same as making $x\vec{u} + y\vec{v}$ closest to $\vec{b}$, i.e., is the same as finding x and y so that $x\vec{u} + y\vec{v}$ is the projection of $\vec{b}$ into the subspace spanned by $\vec{u}$ and $\vec{v}$. So this point of view leads to the same x and y as the first point of view.

least squares solution to an inconsistent system of equations

Here is what emerged from example 1.

Solving the equation $A\vec{x} = \vec{b}$ amounts to trying to write $\vec{b}$ as a combination of the cols of A. When it can't be done (i.e., when the system is inconsistent) then you can still find the combination of cols of A that is *closest* to $\vec{b}$ by projecting $\vec{b}$ onto the col space of A.

It turned out that you can get the normal equations for finding the projection by left multiplying by A^T on each side of the original inconsistent matrix equation

So to find the best *non*-solution to an inconsistent system $$A\vec{x} = \vec{b},$$ left multiply on both sides by A^T to get the normal equations $$A^T A\vec{x} = A^T\vec{b}$$ and solve them instead.

The result is called the *least squares* solution of the original system because from another point of view, it minimizes a sum of square errors.

warning

After you solve normal equations you must know what to do with the solution. This depends on why you were using normal equations in the first place.

If example 1 had just said "find the projection of b onto the subspace of R^3 with basis u,v" then the answer would have been $b_{proj} = -\frac{5}{8}u + \frac{1}{2}v$ (this expresses b_{proj} w.r.t. basis u,v for the subspace) and also $b_{proj} = (-\frac{1}{8}, \frac{1}{4}, -\frac{3}{4})$ (this expresses b_{proj} w.r.t. the basis i,j,k in R^3).

If example 1 had said "find the combination of u and v that is closest to b" then the answer would have been $-\frac{5}{8}u + \frac{1}{2}v$. The answer $(-\frac{1}{8}, \frac{1}{4}, -\frac{3}{4})$ is not good since this is not expressed as a combination of u and v.

But example 1 actually said "find the best non-solution to the inconsistent system in (1)". The answer is *not* $-\frac{5}{8}u + \frac{1}{2}v$ and it is *not* $(-\frac{1}{8}, \frac{1}{4}, -\frac{3}{4})$. The answer is $x = -\frac{5}{8}$, $y = \frac{1}{2}$.

example 2

Look at the system of equations

$$\begin{aligned} x + 2y + z &= 7 \\ 2x \qquad + z &= 2 \\ 6y + z &= 1 \\ -x \qquad + z &= 0 \\ x + y + z &= 0 \end{aligned} \tag{6}$$

It has no solutions. (To satisfy the second and fourth equations you need $x = 2/3$, $z = 2/3$. Then to satisfy the third equation you need $y = 1/18$. But these values don't satisfy the last equation.)

Find the least squares solution.

solution The system of equations in matrix form is

$$\underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 0 & 6 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 7 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{\vec{b}}$$

Then $A^TA = \begin{bmatrix} 7 & 3 & 3 \\ 3 & 41 & 9 \\ 3 & 9 & 5 \end{bmatrix}$, $A^T\vec{b} \begin{bmatrix} 11 \\ 20 \\ 10 \end{bmatrix}$ and the normal equations are

$$A^TA\vec{x} = A^T\vec{b}$$

which amounts to

$$\begin{aligned} 7x + 3y + 3z &= 11 \\ 3x + 41y + 9z &= 20 \\ 3x + 9y + 5z &= 10 \end{aligned}$$

The solution to the normal equations is

$$x = 23/22, \quad y = 2/11, \quad z = 23/22 \tag{7}$$

This is the least squares solution to the inconsistent system in (6).

example 2 continued

Explain what projection was involved in example 2 and why.

solution Trying to solve the system in (6) for x,y,z is equivalent to trying to write $\vec{b}$ as a combination of the cols of A, namely, as x col 1 + y col 2 + z col 3. Since this can't be done, the next best thing is to find x,y,z so that x col 1 + y col 2 + z col 3 is as *close* to $\vec{b}$ as possible. So I found x,y,z by projecting $\vec{b}$ into the subspace spanned by the cols of A.

example 2 continued some more

The system in (6) has no solution. So for any x,y,z that you may pretend is a solution, there are errors involved. What are the errors. And in particular, what is the significance of the least squares solution.

solution For any would-be solution x,y,z, the errors are

$$\begin{aligned}
\text{error1} &= x + 2y + z - 7 \\
\text{error2} &= 2x + z - 2 \\
\text{error3} &= 6y + z - 1 \\
\text{error4} &= -x + z - 0 \\
\text{error5} &= x + y + z - 0
\end{aligned}$$

It is possible to make some of the errors zero but there are no values of x,y,z that make *all* the errors zero because the system of equations had no solution, If you use the least squares solution $x = 23/22$, $y = 2/11$, $z = 23/22$ then at least the sum of the squares of the five errors is minimum.

the least squares line of best fit (an instance in which an inconsistent system of equations turns up)

Suppose an experiment produces the four data points

$$P_1 = (-1,0), \quad P_2 = (0,1), \quad P_3 = (1,3), \quad P_4 = (2,9) \tag{8}$$

and you hope that they will lie on a line with equation of the form

$$y = mx + b. \tag{9}$$

But the points are not collinear; when the data is substituted into (9), you get the inconsistent system of equations

$$\begin{aligned}
-m + b &= 0 \\
0m + b &= 1 \\
m + b &= 3 \\
2m + b &= 9
\end{aligned} \tag{10}$$

There is no solution for m and b. But you can find the least squares solution instead. Let

$$A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$$

Then

$$A^TA = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix}, \quad A^T \begin{bmatrix} 0 \\ 1 \\ 3 \\ 9 \end{bmatrix} = \begin{bmatrix} 21 \\ 13 \end{bmatrix}$$

and the normal equations are

$$\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 21 \\ 13 \end{bmatrix}$$

$$6m + 2b = 21$$
$$2m + 4b = 13$$

The solution is

$$m = \frac{29}{10}, \quad b = \frac{18}{10}.$$

So the line that best fits the data is the line

$$y = \frac{29}{10}x + \frac{18}{10}. \tag{11}$$

warning

After you solve normal equations you must know what to do with the solution. In this fitting-data problem, you ended up projecting (0,1,3,9) onto the subspace spanned by u = (−1,0,1,2) and v = (1,1,1,1), the columns of A. The projection turned out to be $\frac{29}{10}$ u + $\frac{18}{10}$ v. But the answer to the problem is *not* $\frac{29}{10}$ u + $\frac{18}{10}$ v and it is *not* $m = \frac{29}{10}$, $b = \frac{18}{10}$. The answer is the line $y = \frac{29}{10}x + \frac{18}{10}$.

geometric interpretation of the line of best fit

Fig 1 shows the four data points from (8) and a typical line y = mx + b. When you plug an m and b into the inconsistent system in (9), the errors are

error1 = −m+b−0
error2 = 0m+b−1
error3 = m+b−3
error4 = 2m+b−9

Look at error 4 for instance.
2m+b is the y-coord of point D on the line.
9 is the y-coord of point P_4.

The absolute value of error4 is the indicated vertical distance in Fig 1 between P_4 and the line. The same idea holds for the other errors: the absolute value of an error is the difference between the y-coord of a data point and the y-coord of the point on the line y = mx+b above or below the data point.

The line in (10) is the one for which $(\text{error1})^2 + (\text{error2})^2 + (\text{error3})^2 + (\text{error4})^2$ is minimum.

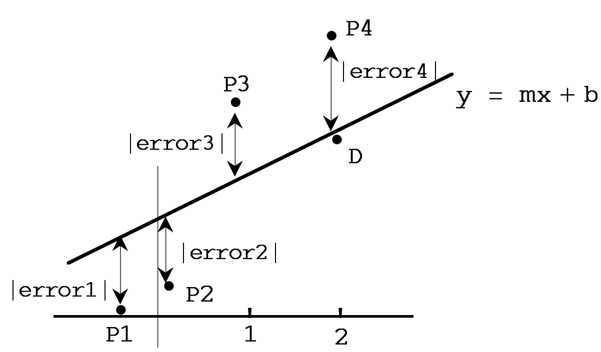


FIG 1

PROBLEMS FOR SECTION 3.3

1. Look at the system of equations $A\begin{bmatrix} x \\ y \end{bmatrix} = \vec{b}$ where $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

(a) Show that the system is inconsistent.
(b) Fill in the blanks: Since the system has no solution, the vector ________ is not in the subspace spanned by the vectors ______________.
(c) Find the least squares solution.
(d) Part (c) involved a projection. What vector was projected, what was it projected onto, and what did the projection turn out to be. And why did you want to find a projection in the first place?
(e) What is the sum of square errors that got minimized and what is that minimum value.

2. Look at the inconsistent system of equations

$$\begin{aligned} x + y + z &= 0 \\ -2x \quad + 2z &= 2 \\ -x \qquad &= 1 \\ 3z &= 0 \\ x + y \quad &= 2 \end{aligned}$$

Find the least squares solution.

3.(a) Use least squares to fit a line to the points
$C = (-1,2)$, $D = (0,0)$, $E = (1,-3)$, $F = (2,-5)$.

(b) Sketch the line, the points C,D,E,F and the errors.
(c) Compute the errors algebraically.
(d) When you found the line of best fit, something got minimized. What was it and what did the minimum value turn out to be.

4. (a) Use least squares to find a formula of the form $y = ax^2 + bx + c$ that best fits this data.

x	0	1	2	3
y	3	2	4	4

(b) Part (a) involved a projection. What vector was projected, what was it projected onto and what did the projection turn out to be.

5. (a) Use least squares to find a formula of the form $z = ax + by$ that best fits this data.

x	2	1	3	−1	1
y	−2	3	1	1	−1
z	0	5	3	−1	1

(b) Find the sum of square errors and its minimum value.

6. Start with a bunch of vectors in R^{32} and a vector v in R^{32} outside the bunch. What does it mean to say u is the vector in the bunch closest to v.

SECTION 3.4 THE GRAM SCHMIDT FORMULAS

Gram Schmidt orthogonalization process

Given a basis $\vec{x_1}, \vec{x_2}, \vec{x_3}, \vec{x_4}, \vec{x_5}$ for a 5-dim subspace V of R^n.

You can get an *orthogonal* basis $\vec{u_1}, \vec{u_2}, \vec{u_3}, \vec{u_4}, \vec{u_5}$ for V with these formulas.

$$\vec{u_1} = \vec{x_1}$$

$$\vec{u_2} = \vec{x_2} - \frac{u_1 \cdot x_2}{u_1 \cdot u_1}\vec{u_1}$$

$$\vec{u_3} = \vec{x_3} - \frac{u_1 \cdot x_3}{u_1 \cdot u_1}\vec{u_1} - \frac{u_2 \cdot x_3}{u_2 \cdot u_2}\vec{u_2}$$

$$\vec{u_4} = \vec{x_4} - \frac{u_1 \cdot x_4}{u_1 \cdot u_1}\vec{u_1} - \frac{u_2 \cdot x_4}{u_2 \cdot u_2}\vec{u_2} - \frac{u_3 \cdot x_4}{u_3 \cdot u_3}\vec{u_3}$$

$$\vec{u_5} = \vec{x_5} - \frac{u_1 \cdot x_5}{u_1 \cdot u_1}\vec{u_1} - \frac{u_2 \cdot x_5}{u_2 \cdot u_2}\vec{u_2} - \frac{u_3 \cdot x_5}{u_3 \cdot u_3}\vec{u_3} - \frac{u_4 \cdot x_5}{u_4 \cdot u_4}\vec{u_4}$$

You can get an ortho*normal* basis by normalizing $\vec{u_1}, \vec{u_2}, \vec{u_3}, \vec{u_4}, \vec{u_5}$.

Another way of looking at the formulas is that they replace a bunch of independent vectors in R^n with an equal number of new vectors that are combinations of the old ones, so that the new ones are orthogonal (and nonzero).

proof

I'm going to choose five new vectors $\vec{u_1}, \vec{u_2}, \vec{u_3}, \vec{u_4}, \vec{u_5}$.

Each new vector will be a combination of $\vec{x_1}, \vec{x_2}, \vec{x_3}, \vec{x_4}, \vec{x_5}$ so that the new ones are *in* V. And I'll choose so that the new ones are orthogonal (and nonzero).

Start by letting $u_1 = x_1$.

Now I want to replace x_2 by something orthog to u_1.

Let $x_{2_{\text{proj}}}$ be the projection of x_2 into the 1-dim subspace with basis u_1 (Fig 1).

Then $x_{2_{\text{proj}}} = \dfrac{u_1 \cdot x_2}{u_1 \cdot u_1} u_1$. The vector $x_2 - x_{2_{\text{proj}}}$ is orthog to u_1. Choose it to be u_2.

(Note that u_2 is a combination ultimately of x_1 and x_2 as advertised.)

Now I have u_1, u_2 orthog and I want to replace x_3 by a vector orthog to u_1 and u_2.

Project x_3 into the 2-dim subspace with orthog basis u_1, u_2 (Fig 2). Then

$$x_{3_{\text{proj}}} = \frac{u_1 \cdot x_3}{u_1 \cdot u_1}\vec{u_1} + \frac{u_2 \cdot x_3}{u_2 \cdot u_2}\vec{u_2}$$

The vector $x_3 - x_{3_{\text{proj}}}$ is orthog to u_1 and u_2. Choose it to be u_3.

Similarly for u_4 and u_5.

Now I have 5 orthogonal vectors $\vec{u_1}, \ldots, \vec{u_5}$ in V, matching the formulas given above.

Here's the check that they are nonzero.

u_1 is nonzero because $u_1 = x_1$ and x_1 is a basis vector which can't be zero..

u_2 is nonzero because x_2 is not a multiple of x_1 so x_2 is not in the 1-dim subspace with basis u_1 so $x_2 \neq x_{2_{\text{proj}}}$.

etc.

So $\vec{u}_1, \ldots, \vec{u}_5$ is a basis for V (5 nonzero orthogonal vectors in a 5-dim subspace are a basis).

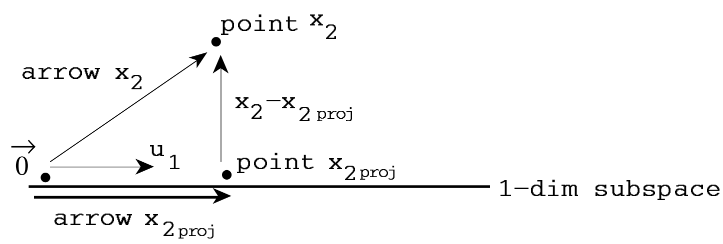


FIG 1

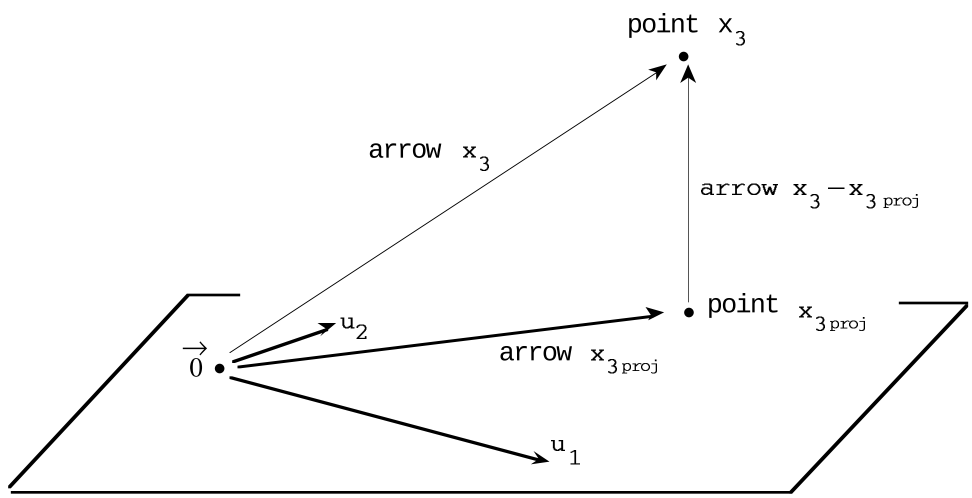


FIG 2

example 1
Let

$$x_1 = (1,1,1,-1)$$
$$x_2 = (2,-1,-1,1)$$
$$x_3 = (-1,2,2,1)$$

I'll find an orthogonal basis for the subspace of R^4 with basis x_1, x_2, x_3 and then find an orthonormal basis for the subspace.

The Gram Schmidt formulas give this orthogonal basis:

$$u_1 = x_1 = (1,1,1,-1)$$
$$u_2 = (2,-1,-1,1) - \frac{-1}{4}(1,1,1,-1) = \left(\frac{9}{4}, -\frac{3}{4}, -\frac{3}{4}, \frac{3}{4}\right)$$
$$u_3 = (-1,2,2,1) - \frac{2}{4}(1,1,1,-1) + \frac{2}{3}\left(\frac{9}{4}, -\frac{3}{4}, -\frac{3}{4}, \frac{3}{4}\right) = (0,1,1,2)$$

An orthonormal basis is

$$u_{1\,unit} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)$$
$$u_{2\,unit} = \left(\frac{3}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}\right)$$
$$u_{3\,unit} = \left(0, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)$$

PROBLEMS FOR SECTION 3.4

1.Look at the subspace with the following basis. Find an orthogonal basis and then an orthonormal basis for the subspace.

(a) $\overrightarrow{x_1} = (0,1,0)$, $\overrightarrow{x_2} = (1,2,3)$

(b) $\vec{p} = (1,2,2)$, $\vec{q} = (1,3,1)$

(c) u = (1,0,0,1), v = (1,0,1,0), w = (2,1,1,1)

2. Start with two independent vectors u and v in R^{34}.
Suppose all you know about them is $u \cdot u = 4$, $v \cdot v = 3$, $u \cdot v = 2$.
(a) u and v themselves are not orthogonal.
Find two combinations of u and v that are orthogonal.
And when you are finished, check that they really are orthogonal.
(b) Now find two combinations of u and v that are orthonormal.

3. Go back to the subspace spanned by the vectors in #1(c).
Let q = (5,6,6,5). Find (fairly easily) the projection of q on the subspace.

4. Suppose u,v,w,p is a basis for R^4. You use the Gram Schmidt process on them. Why is that dumb.

CHAPTER 4 MATRICES CONTINUED

SECTION 4.1 RANK OF A MATRIX

rank list

Given matrix M, the following are equal:

(1) maximal number of ind cols (i.e., dim of the col space of M)
(2) maximal number of ind rows (i.e., dim of the row space of M)
(3) number of cols with pivots in the echelon form of M
(4) number of nonzero rows in the echelon form of M

You know that (1) = (3) and (2) = (4) from Section 3.1.
To see that (3) = (4) just stare at some echelon forms.

This one number (that all four things equal) is called the *rank* of M.
As a special case, a zero matrix is said to have rank 0.

how row ops affect rank

Row ops don't change the rank because they don't change the max number of ind cols or rows.

example 1

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 4 & -2 & 0 \\ 4 & 8 & -4 & 0 \end{bmatrix}$$ has rank 1 (maximally one ind col, by inspection)

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$ has rank 0

example 2

Let $M = \begin{bmatrix} 2 & 5 & 4 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & 1 & -1 & 0 \\ 2 & 11 & 7 & 0 \end{bmatrix}$

To find the rank of M, use row ops

$R3 = R1 + R3$
$R4 = -R1 + R4$
$R2 \leftrightarrow R3$
$R4 = -R2 + R4$

to get the unreduced echelon form $\begin{bmatrix} 2 & 5 & 4 & 0 \\ 0 & 6 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

Cols 1,2,4 have pivots. So the rank of M is 3.

how the rank is limited by the size of the matrix

If A is 7×4 then its rank is either 0 (if it's the zero matrix), 1, 2, 3 or 4. The rank can't be 5 or larger because there can't be 5 ind cols when there are only 4 cols to begin with.

example 3

If M is 7×6 and has rank 4 then the 6 cols are dep, any set of 5 cols is dep, some foursome of cols is ind. Similarly, the 7 rows are dep, any 6 rows are dep, any 5 rows are dep, some set of 4 rows is ind.

example 4

If A is 6 × 7 and the 6 rows are dependent then the rank is 0,1,2,3,4 or 5.

invertible rule

Let M be an n × n matrix. The following are equivalent; i.e., either all are true or all are false.

(1) M is invertible (nonsingular).
(2) $|M| \neq 0$.
(3) Echelon form of M is I.
(4) Rows of M are independent.
(5) Cols of M are independent.
(6) Rank of M is n.

proof that (6) belongs on the list

Item (6) is equivalent to (5) because for an n × n matrix, ind cols means maximally n ind cols which means the rank is n.

order of a determinant

If a matrix is n × n then its determinant is said to have order n.

subdeterminants

The fat dots in Fig 1 illustrate the idea of 3 × 3 submatrix. The determinant of the fat dots is a subdeterminant of order 3 as opposed to the original det which is of order 5.

A 5 × 5 matrix has lots of subdets of orders 4,3,2,1 and has one subdet (the whole thing) of order 5.

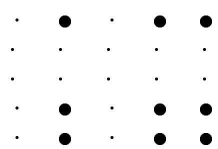


FIG 1

rank list

Given a matrix M. The following are all equal and are called the rank of M:

(1) maximal number of ind cols (i.e., dim of the col space of M)
(2) maximal number of ind rows (i.e., dim of the row space of M)
(3) number of cols with pivots in the echelon form of M
(4) number of nonzero rows in the echelon form of M
(5) largest order of all the nonzero subdeterminants of M

The proof that (5) belongs on the list is omitted because it gets messy.

example 5

If a 5 × 7 matrix has rank 3 then every 5 × 5 subdeterminant is 0, every 4 × 4 subdeterminant is 0 and at least one (but not necessarily all) of the 3 × 3 subdeterminants is not 0.

PROBLEMS FOR SECTION 4.1

1. Find the rank

(a) $\begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 & 5 \\ 0 & 1 & 3 \\ 2 & 5 & 19 \\ -1 & 1 & 1 \\ 3 & 4 & 18 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & 6 \end{bmatrix}$

2. Suppose M has 9 cols and its rank is 5. Decide, if possible, whether the following are ind or dep.
(a) first six cols of M
(b) first five cols of M
(c) first four cols of M

3. Suppose M is 7×9 (7 rows, 9 cols). What can you conclude about the rank of M if
(a) that's all the information you have
(b) row 3 = 2 row 7 + 8 row 4
(c) first five rows are ind
(d) first 5 rows are ind and every set of 6 rows is dep.
(e) first 7 cols are ind
(f) rows 1 and 2 are ind, rows 1,2,3 are dep

4. Suppose the echelon form of M is

$$\begin{matrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix}$$

What can you conclude about $|M|$ and about subdeterminants of M.

5. Suppose rank M = 5. Find rank M^T.

6. Suppose rank A = 6 and A is invertible. Find the rank of A^{-1}.

7. If rank B = 2 find rank(3B).

8. Let

$$M = \begin{bmatrix} t & \sqrt{2} & 0 \\ \sqrt{2} & t & \sqrt{2} \\ 0 & \sqrt{2} & t \end{bmatrix}$$

Find the rank of M (it will depend on t).

9. M is 5×5. What can you conclude about the rank of M if
(a) one of its 3×3 subdeterminants is 6
(b) the 3×3 subdeterminant in the NE corner is 0 and the 3×3 subdet in the SW corner is 8
(c) one of its 3×3 subdets is 0, another of its 3×3 subdets is 8 and all the 4×4 subdets are 0
(d) all the 3×3 subdets are 0
(e) one of its 3×3 subdets is 0

SECTION 4.2 ORTHOGONAL MATRICES

orthogonal matrix rule

Let M be a (real) $n \times n$ matrix. The following are equivalent (i.e., if any one of the three holds then they all hold, if any one fails then they all fail).

(1) $M^{-1} = M^T$, i.e., $MM^T = I$; i.e. $M^TM = I$.

(2) Cols of M are orthonormal vectors in R^n (orthogonal *and* unit length).

(3) Rows of M are orthonormal vectors in R^n.

A matrix with these properties is called *orthogonal*.

Note that, for a square matrix, this rule says all of the following (and then some):
If the rows are orthonormal then the cols are also orthonormal.
If the cols are orthonormal then the rows are also orthonormal.
If the rows are orthonormal then the inverse of the matrix is easy to find; it's the transpose.
If the cols are orthonormal then the inverse of the matrix is easy to find; it's the transpose.

warning

Items (2) and (3) in the orthogonal matrix rule involve ortho*normal* rows and cols, not just orthogonal rows and cols.

proof that (1) is equivalent to (3)

Suppose M is 3×3. Let u,v,w be the rows of M so that u,v,w are the cols of M^T. then

$$MM^T = \begin{bmatrix} u\cdot u & u\cdot v & u\cdot w \\ v\cdot u & v\cdot v & v\cdot w \\ w\cdot u & w\cdot v & w\cdot w \end{bmatrix}$$

So

$$\begin{aligned} MM^T = I \quad &\text{iff } u\cdot u = v\cdot v = w\cdot w = 1 \text{ and } u\cdot v = u\cdot w = v\cdot w = 0 \\ &\text{iff } u,v,w \text{ are orthog and } \|u\| = \|v\| = \|w\| = 1 \\ &\text{iff } u,v,w \text{ are orthonormal} \end{aligned}$$

proof that (1) is equivalent to (2)

This can be done similarly by letting p,q,r be the cols of M and considering M^TM instead of MM^T.

example 1

Let

$$M = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ -2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 1/\sqrt{30} & -5\sqrt{30} & 2/\sqrt{30} \end{bmatrix}$$

By inspection, the rows are unit vectors (i.e., $\|\text{row } 1\| = \|\text{row } 2\| = \|\text{row } 3\| = 1$. And the rows are orthogonal (row 1·row 2 = row 1·row 3 = row 2·row 3 = 0). So M is an orthogonal matrix. Its cols are orthonormal and $M^{-1} = M^T$.

the basis changing matrix

If a new basis for R^n is orthonormal then the basis changing matrix P has orthonormal cols so it is an orthogonal matrix and easy to invert; namely $P^{-1} = P^T$.

For example, if the axes in R^2 are rotated by 32° then the new orthonormal basis vectors are

$$u = (\cos 32^\circ, \sin 32^\circ)$$
$$v = (-\sin 32^\circ, \cos 32^\circ) \; .$$

And

$$P = \begin{bmatrix} \cos 32^o & -\sin 32^o \\ \sin 32^o & \cos 32^o \end{bmatrix}$$

$$P^{-1} = P^T = \begin{bmatrix} \cos 32^o & \sin 32^o \\ -\sin 32^o & \cos 32^o \end{bmatrix}$$

determinant of an orthogonal matrix

(1) If M is orthogonal then $|M| = \pm 1$.
This also means that if $|M| \neq \pm 1$ then M can't be orthogonal.

(2) But orthogonal matrixes are not the only matrices whose dets are ± 1; in other words, if $|M| = \pm 1$ then M is not necessarily an orthogonal matrix.

proof of (1)

Suppose A is orthogonal. Then

$AA^T = I$

$|AA^T| = 1$

$|A||A^T| = 1$

$|A|^2 = 1$

$|A| = \pm 1$

proof of (2)

Here's a counterexample. Let $M = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}$.

Then det M = 1 but M is not orthogonal.

warning repeated

A matrix with merely orthogonal cols is a *nothing*.
An orthogonal matrix has to have ortho*normal* cols.

how to show that a matrix is orthogonal

method 1 Look at its rows (or cols) to see if they are orthonormal.
method 1 Multiply the matrix by its transpose to see if the product is I.

how to show that a matrix is not orthogonal

method 1 Look at its rows (or cols) to see that they are not orthonormal.
method 2 Or show that the matrix times its transpose is not I.
method 3 Find the det. If the det is not ± 1 then the matrix is not orthog. (But if the det is ± 1 then you have no conclusion.)

example 2

Show that if M is orthogonal then $-M$ is also orthogonal.
solution method 1
If M is orthogonal then its col vectors are orthonormal. Intuitively, multiplying by -1 reverses the direction of an arrow but doesn't change it length, so the cols of $-M$ are still orthogonal and still have unit length. So $-M$ is orthog.

For this to count as a legal argument in R^n you have to show that in R^n, if u and v are orthogonal then $-u$ and $-v$ are also orthog. Here's how:

If $u \cdot v = 0$ then

$$\begin{aligned} (-u)\cdot(-v) &= (-1)(-1)(u\cdot v) && \text{by dot product rules)} \\ &= 0 && \text{because u and v are orthogonal} \end{aligned}$$

And you have to show that if u is a unit vector then $-u$ is also a unit vector.
Here's how:

$\|-u\| = |-1|\,\|u\|$ by norm rules
$= \|u\|$ because the abs value of -1 is 1
$= 1$ because u is a unit vector QED

method 2 (neater)

I'll show that $(-M)(-M)^T = I$:

$$
\begin{aligned}
(-M)(-M)^T &= (-M)(-M^T) && \text{T rule} \\
&= (-1)(-1)MM^T && \text{matrix algebra} \\
&= MM^T \\
&= I && \text{since M is orthog}
\end{aligned}
$$

PROBLEMS FOR SECTION 4.2

1. Are the following matrices orthogonal.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 2 & 2 \\ 2 & -2 \end{bmatrix} \qquad E = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

2. Suppose M is a square matrix.
(a) Does it count as an orthogonal matrix if someone tells you that its rows are orthonormal but tells you nothing about its columns.
(b) Does it count as an orthogonal matrix if someone tells you its rows are orthonormal and its columns aren't ???????????
(c) Does it count as an orthogonal matrix if its rows are orthogonal.

3. For each matrix, find all values of a and b that make M orthogonal.

(a) $M = \begin{bmatrix} a & 2 \\ b & 3 \end{bmatrix}$ (b) $M = \begin{bmatrix} a & \frac{1}{\sqrt{10}} \\ b & \frac{3}{\sqrt{10}} \end{bmatrix}$.

4. Look at this statement:

If A is square then A and A^T commute (meaning $AA^T = A^TA$).

(a) Show that it is *not* true in general
(b) Show it *is* true in the special case that A is orthogonal.

5. Find all 3×3 diagonal orthogonal matrices.

6. Let $A = \begin{bmatrix} 1 & 2 \\ 6 & -3 \end{bmatrix}$.

(a) Then A is a square matrix with orthogonal rows and non-orthogonal columns. Does this violate the orthogonal matrix rule.
(b) Make a new matrix B by normalizing the rows in A. What would you expect of the *columns* in B and check that it really happens.

7. (a) Find a counterexample to disprove that that the sum of orthog matrices is orthog.
 (b) Show that the product of orthogonal matrices is orthog.
 (c) Show that the transpose of an orthog matrix is orthog.
 (d) Show that the inverse of an orthog matrix is orthog.

8. Suppose M is orthogonal. Will 5M also be orthog?

9. Let

$$M = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Find M^{-1} easily.

10. Let $u = (u_1,u_2,u_3)$, $v = (v_1,v_2,v_3)$, $w = (w_1,w_2,w_3)$ be a new orthonormal basis for R^3. Let x,y,z be the old coordinates of a point and let X,Y,Z be its new coordinates w.r.t. basis u,v,w.
(a) Express x,y,z in terms of X,Y,Z.
(b) Express the X,Y,Z in terms of x,y,z (easily).

11. Suppose M is 3×3 with cols u,v,w where u,v,w are orthogonal and $\|u\| = \|v\| = \|w\| = 2$.
(a) Express M^{-1} in terms of M^T.
(b) What can you conclude about $|M|$.

12. Suppose M is 3×3 with cols u,v,w where u,v,w are orthogonal and $\|u\| = 2$, $\|v\| = 3$, $\|w\| = 4$.
What can you conclude about $|M|$.

13. Let $M = \begin{bmatrix} 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix}$

The cols are orthonormal but the rows aren't. I thought that if the cols were orthonormal then the rows must also be orthonormal. So what happened.

REVIEW PROBLEMS FOR CHAPTERS 3 AND 4

1. Let p = (2,1,4,2,0), q = (1,0,1,0,1). Find an orthonormal basis for the subspace of R^5 with basis p,q.

2. Find all values of k that make (1,1,1), (1,2,4), $(1,k,k^2)$ a basis for R^3.

3. Show that if S is symmetric and B is orthogonal then $B^{-1}SB$ is symmetric.

4. Let

u = (1,0,2)
v = (2,1,3)
w = (0,1,1)

Show that u,v,w is a basis for R^3 and express the vector $a\vec{i} + b\vec{j} + c\vec{k}$ in terms of them (a) using row ops (b) using the basis changing matrix

5. Show that if M is orthogonal then the operator M preserves norms in the sense that the norm of $M\vec{u}$ is the same as the norm of $\vec{u}$.
Use the fact that if u and v are col vectors then the dot product u·v can be written as the matrix product u^Tv (end of Section 2.1).

6. Look at the subspace of R^6 spanned by u,v,w,q.
I lined up u,v,w, q as rows and then again as columns and got the two echelon forms respectively.

$$\text{echelon form of } [u\ v\ w\ q] = \begin{matrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}$$

(u,v,w,q were the original cols)

$$\text{echelon form of} \begin{bmatrix} u \\ v \\ w \\ q \end{bmatrix} = \begin{matrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

(u,v,w,q were the original rows)

(a) Find as many bases as you can for the subspace by choosing from among the spanning vectors u,v,w,q.
(b) Find the standard basis for the subspace.
(c) Describe the nice pattern that characterizes the members of the subspace.
(d) Find ten more bases for the subspace without taking the easy way out and just using multiples of basis vectors you already have.

7. Find the line that best fits the data points P = (1,0), Q = (2,2), R = (3,7).

8. All vectors in this problem are in R^4.
If you had a computer that could do nothing but find echelon forms, how would you use it to answer each of the following.
(a) Are vectors u,v,w,p independent.
(b) Is y in the subspace spanned by u,v,w.
(c) Is a 4×4 matrix A invertible.
(d) Find the rank of a matrix M.

CHAPTER 5 SYSTEMS OF EQUATIONS

SECTION 5.1 GAUSSIAN ELIMINATION

matrix form of a system of equations

The system

$$\begin{aligned} 2x + 3y + 4z &= 1 \\ 5x + 6y + 7z &= 2 \end{aligned}$$

can be written as

$$A\vec{x} = \vec{b}$$

where

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

The system is abbreviated by writing

$$(1) \qquad \left.\begin{array}{ccc|c} 2 & 3 & 4 & 1 \\ 5 & 6 & 7 & 2 \end{array}\right.$$

The matrix A is called the *coefficient matrix*. The 2×4 matrix in (1) is called the *augmented matrix* and is denoted $A|b$.

Gaussian elimination

Row ops on $A|b$ amount to interchanging two equations or multiplying an equation by a nonzero constant or adding a multiple of one equation to another. They do not change the solution so they may be used to simplify the system. In particular, performing row ops on $A|b$ until A is in echelon form is called *Gaussian elimination*.

There are two possibilities (Fig 1).

1. The row ops produce a row of the form

$$(2) \qquad 0 \quad 0 \quad 0 \quad 0 \;|\text{nonzero}$$

Then the system has no solution and is called *inconsistent*.

For example, if a system row ops to

$$\left.\begin{array}{cccc|c} 1 & 0 & 2 & 4 & 2 \\ 0 & 1 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{array}\right.$$

then it has no solutions because the third row is the equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 6$$

which is impossible to satisfy.

2. The row ops do not produce a row of the form $0\ 0\ 0\ 0|\text{nonzero}$.
   Then the system has a solution (at least one) and is called *consistent*.
   There are two subcases (see the summary in the box below).

   2a. All the echelon cols of A have pivots.
   Then there is exactly one solution.
   For example if the row ops produce

$$\left.\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right.$$

and the variables are named x,y,z then the solution is x=2, y=5, z=6.

2b. Not all the echelon cols of A have pivots.
Then there are infinitely many solutions.
It is always possible to solve for some of the variables in terms of the others. Those others are called *free variables* or *parameters*.
One way to do this is to *choose the variables corresponding to the cols without pivots to be free. If the system $Ax = b$ has n variables* (so that A has n cols) *and* $\operatorname{rank} A = r$ (so that there are r cols with pivots in the echelon form of A) *then the solution has $n - r$ free variables.*

For example, if the row ops produce

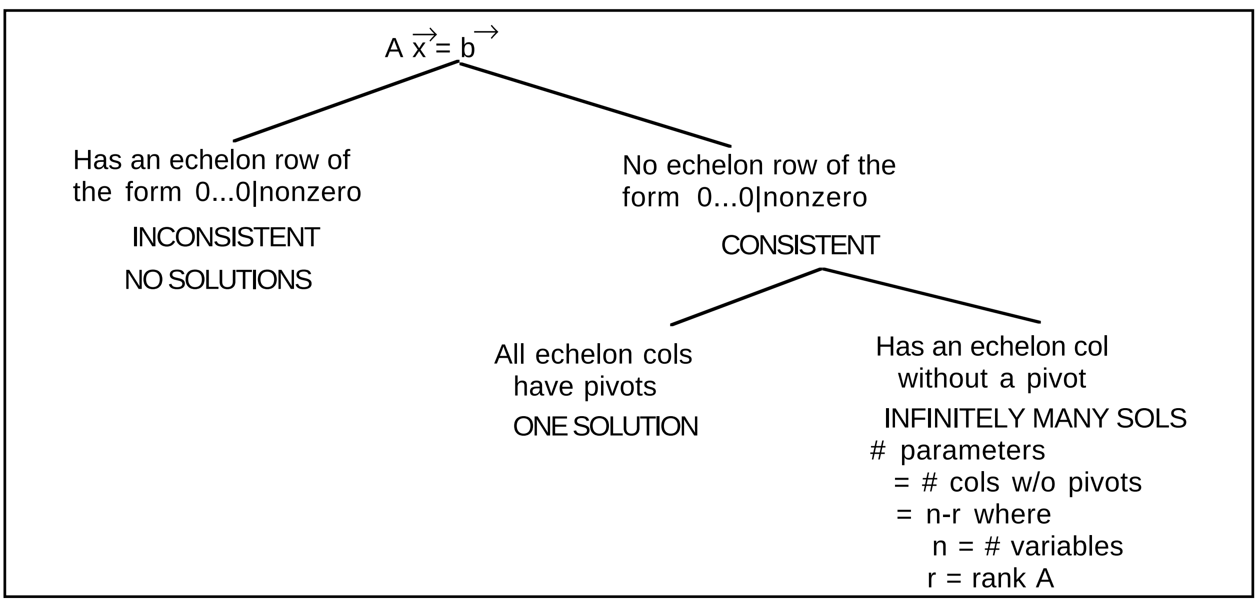
$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 3 & 5 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

and the variables are named x,y,z,w then one way to write the solution is

$$\begin{aligned} x &= 5 - 3w - 2y \\ z &= 6 - 4w \qquad (w,\ y \text{ free}) \end{aligned}$$

Another way to write the sol is

$$\begin{aligned} x &= 5 - 3s - 2t \\ y &= t \\ z &= 6 - 4s \\ w &= s \end{aligned}$$



example 1

Solve

$$\begin{aligned} 2x_1 + 4x_2 - 2x_3 + 8x_4 + 4x_5 &= 6 \\ 3x_1 + 6x_2 + x_3 + 12x_4 - 2x_5 &= 1 \\ 9x_1 + 18x_2 + x_3 + 36x_4 + 38x_5 &= 0 \end{aligned}$$

solution Begin with

$$\left[\begin{array}{ccccc|c} 2 & 4 & -2 & 8 & 4 & 6 \\ 3 & 6 & 1 & 12 & -2 & 1 \\ 9 & 18 & 1 & 36 & 38 & 0 \end{array}\right]$$

and do row ops

$$R1 = \frac{1}{2}R1$$
$$R2 = -3R_1 + R2$$
$$R3 = -9R1 + R3$$
$$R2 = \frac{1}{4}R2$$
$$R1 = R2 + R1$$
$$R3 = -10R2 + R3$$
$$R3 = \frac{1}{40}R3$$
$$R2 = 2R3 + R2$$

to get

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 4 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -94/40 \\ 0 & 0 & 0 & 0 & 1 & -7/40 \end{array}\right]$$

Choose x_2 and x_4, the variables corresponding to the cols without pivots, to be the free variables (other choices are possible but I like these) and solve for x_1, x_3, x_5 in terms of them. A final sol is

$$x_5 = -\frac{7}{40}$$
$$x_3 = -\frac{94}{40}$$
$$x_1 = 1 - 4x_4 - 2x_2$$
$$(x_2,\ x_4 \text{ free})$$

Another version of the sol is

$$x_5 = -\frac{7}{40}$$
$$x_4 = t$$
$$x_3 = -\frac{94}{40}$$
$$x_2 = s$$
$$x_1 = 1 - 4t - 2s$$

example 2

Suppose A is 4×3 and $Ax = b$ has infinitely many solutions. Consider the new system $Ax = c$. Will it have infinitely many solutions also; i.e., what happens if you change the righthand side of the system.

solution Since $Ax = b$ has infinitely many sols, the echelon form of A must have at least one col without a pivot and $A|b$ must row op into something like this:

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 8 \\ 0 & 1 & 7 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{or} \quad \left[\begin{array}{ccc|c} 1 & 3 & 0 & 8 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{or} \quad \left[\begin{array}{ccc|c} 1 & 3 & 7 & 8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

warning I'm not saying that the *original* system looks like this. But $A|b$ must *row operate* into something like this.

Now look at $Ax = c$, same A but different righthand side.
Can't have just one sol because the echelon form of A has at least one col without a pivot.
Can have infinitely many sols or no solutions: If $A|b$ row opped to

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 8 \\ 0 & 1 & 7 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

it's possible for $A|c$ to row op to say

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 5 \\ 0 & 1 & 7 & 5 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right.$$

and have no solutions, and it's also possible for $A|c$ to row op to

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & \pi \\ 0 & 1 & 7 & e \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right.$$

and have infinitely many sols.

So if $Ax = b$ has infinitely many sols then $Ax = c$ has either infinitely many sols or no solutions (but can't have just one solution).

solving with unreduced echelon form and back substitution (much more efficient)

Row operate on the system so that the coeff matrix is in unreduced echelon form (upper triangular form). Then starting with the *last* row, solve for the first variable in each row and back substitute as you go along.

For example, if the row operations produce

$$\left[\begin{array}{ccccc|c} 1 & 2 & 3 & 5 & 7 & 2 \\ 0 & 0 & 4 & 6 & 8 & 6 \\ 0 & 0 & 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right.$$

then

$$x_5 = 3 \qquad \text{from the third row}$$

$$x_3 = \frac{6 - 8x_5 - 6x_4}{4} \qquad \text{from second row}$$

$$= \frac{6 - 8(3) - 6x_4}{4} \qquad \text{back substitute}$$

$$= \frac{-18 - 6x_4}{4}$$

$$x_1 = 2 - 7x_5 - 5x_4 - 3x_3 - 2x_2 \qquad \text{from first row}$$

$$= 2 - 7(3) - 5x_4 - 3\left(\frac{-18-6x_4}{4}\right) - 2x_2 \qquad \text{back substitute}$$

$$= -\frac{11}{2} - 2x_2 - \frac{1}{2}x_4$$

The solution can also be written as

$$x_5 = 3$$

$$x_4 = t$$

$$x_3 = \frac{-18 - 6t}{4}$$

$$x_2 = s$$

$$x_1 = -\frac{11}{2} - 2s - \frac{1}{2}t$$

PROBLEMS FOR SECTION 5.1

1. Here are some systems of equations, already in echelon form.
Let the variables be named $x_1, x_2, \ldots$.
Solve and identify the free variables. Then express the solution using parameters r,s,t,...

(a) $\left[\begin{array}{cccccc|c} 1 & 2 & 0 & 3 & 5 & 0 & 5 \\ 0 & 0 & 1 & 4 & 6 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$ (b) $\left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 6 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{array}\right]$

(c) $\left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$ (d) $\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$ (e) $\left[\begin{array}{cccc|c} 1 & 2 & 0 & 0 & 5 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

2. Let
$u = (1,0,1,1)$
$v = (0,1,1,0)$
$w = (1,1,2,1)$
$y = (2,-1,1,2)$

Decide if y is in the subspace spanned by u,v,w using
(a) ideas from Section 3.1
(b) ideas from this section

If it is, express y as a combination of u,v,w.

3. Look at the system

$$\begin{aligned} x - 2y + 3z &= 1 \\ 2x + ky + 6z &= 6 \\ -x + 3y + (k-3)z &= 0 \end{aligned}$$

For what values of k will it have
(a) no solutions (b) one solution (c) infinitely many solutions

4. True or False. If False, what would the correct conclusion be.
(a) If there are fewer equations than unknowns then the unknowns are "underdetermined" and there will be infinitely many sols.
(b) If there are more equations than unknowns then the unknowns are "overdetermined" and there are no solutions.
(c) If there are the same number of unknowns as equations then everything is hunky-dory and there will be exactly one solution.

5. Suppose A is 5×7 and the 2×2 subdet in the northeast corner is nonzero but all the 3×3 subdets are 0. What can you conclude about the number of solutions and number of free variables for the system $A\vec{x} = \vec{b}$.

6. Solve using row ops and back substitution.

(a) $\left[\begin{array}{ccccc|c} 2 & 3 & 4 & 6 & 1 & 4 \\ 0 & 0 & 5 & 7 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(b)
$$\begin{aligned} 2x + 8y + 5z + 2u - 6v &= 8 \\ 4y + 6z + 3u + 3v &= 9 \\ 2x + 12y + 11z + 5u - 3v &= 17 \\ 4y + 6z + 3u + 6v &= 9 \end{aligned}$$

(c)
$$\begin{aligned} 2x + y + z &= 1 \\ 4x + y &= -2 \\ -2x + 2y + z &= 7 \end{aligned}$$

7. A system of equations in unknowns x,y,z,w row ops to

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 3 & 5 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Solve and, if possible, make the free variables
(a) w and y
(b) x and w
(c) y and z
(d) x and y

8. Suppose A is 4×3. If $A\vec{x} = \vec{b}$ has one solution, what can you conclude about solutions to $A\vec{x} = \vec{c}$.

9. What can you conclude about sols to $A\vec{x} = \vec{b}$ if
(a) A is 3×5 with rank 3
(b) A is 3×5 with rank 2
(c) A is 5×3 with rank 3

10. Let

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

For what vectors $\vec{b}$ is $M\vec{x} = \vec{b}$ consistent. In that consistent case, solve.

11. You have a fixed 3×4 matrix A (3 rows, 4 columns) with rank 3.
You're looking for a right inverse for A; i.e., you're looking for a matrix

$$B = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ x_4 & y_4 & z_4 \end{bmatrix}$$

so that AB = I.
Can you find such a B. If so, how many.
Suggestion: Think about systems of equations and remember that A is fixed and B is filled with unknowns.

SECTION 5.2 HOMOGENEOUS SYSTEMS

homog systems and the null space of a matrix

A homog system is one of the form $M\vec{x} = \vec{0}$.
It is always consistent since it at least has the (trivial) solution

$$x_1 = 0,\ x_2 = 0,\ \ldots,\ x_n = 0.$$

> A homog system will always have either one solution (the trivial one) or infinitely many sols (the trivial one plus infinitely many others).
>
> Furthermore, the set of solutions to $M\vec{x} = \vec{0}$ is a subspace called the *null space* of M.
> If M has n cols (so that the system has n variables) and rank M = r then *the null space is an (n – r)-dim subspace of* R^n; the dimension of the null space is the number of free variables.

The dimension of the null space of M is called the *nullity* of M.

At worst, the null space is a 0-dim subspace containing only $\vec{0}$.
Here's an example to illustrate why the sols to a homog system are a subspace and to show how to find a basis for the null space.

Suppose $M|\vec{0}$ row ops to

$$\left[\begin{array}{cccccc|c} 1 & -7 & 0 & 2 & 5 & 0 & 0 \\ 0 & 0 & 1 & -3 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Then

$$\begin{aligned} x_6 &= 0 \\ x_3 &= 6x_5 + 3x_4 \qquad (1) \\ x_1 &= -5x_5 - 2x_4 + 7x_2 \end{aligned}$$

The solution can also be written as

$$\begin{aligned} x_1 &= -5r - 2s + 7t \\ x_2 &= t \\ x_3 &= 6r + 3s \\ x_4 &= s \\ x_5 &= r \\ x_6 &= 0 \end{aligned}$$

In vector notation, the solution is

$$(2) \qquad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r \underbrace{\begin{bmatrix} -5 \\ \boxed{0} \\ 6 \\ \boxed{0} \\ \boxed{1} \\ 0 \end{bmatrix}}_{\vec{u}} + s \underbrace{\begin{bmatrix} -2 \\ \boxed{0} \\ 3 \\ \boxed{1} \\ \boxed{0} \\ 0 \end{bmatrix}}_{\vec{v}} + t \underbrace{\begin{bmatrix} 7 \\ \boxed{1} \\ 0 \\ \boxed{0} \\ \boxed{0} \\ 0 \end{bmatrix}}_{\vec{w}}$$

The set of sols is the set of all combinations of u,v,w so it is a subspace of R^6. Furthermore u,v,w are ind (look at their 2nd, 4th and 5th components to see that no one can be a combination of the other). So the sols are a 3-dim subspace with basis u,v,w. The dimension of the subspace matches the number of free variables.

Now that you know the sols are a subspace, here's another way to extract a basis from the solution in (1) without rewriting it as (2). Assign values to the three free variables so as to get 3 ind solutions. The easiest way to do this is to let

$$x_2 = 1,\ x_4 = 0,\ x_5 = 0$$

to get solution (7,1,0,0,0,0); then let

$$x_2 = 0,\ x_4 = 1,\ x_5 = 0$$

to get solution (−2,0,3,1,0,0); then let

$$x_2 = 0,\ x_4 = 0,\ x_5 = 1$$

to get solution (−5,0,6,0,1,0).

homogeneous versus non-homogeneous

The set of solutions to a *homogeneous* system of equations is a subspace but the set of solutions to a *non*-homogeneous system of equations can never be a subspace. For one thing it never contains $\vec{0}$ because a *non*-homogeneous system never has the solution $x_1 = 0,\ x_2 = 0, \ldots, x_n = 0$.

example 1

Suppose the system $M\vec{x} = \vec{0}$ row ops to

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Then

$$x_3 = -4x_4$$

$$x_1 = -3x_4 - 2x_2 \tag{3}$$

The sols are a 2-dim subspace of R^4 since there are 2 free variables.

To pick a basis let $x_4 = 1,\ x_2 = 0$ in (3) to get solution u = (−3,0,−4,1); and set $x_4 = 0,\ x_2 = 1$ in (3) to get solution v = (−2,1,0,0). Then u and v are two ind solutions and are a basis for the subspace of solutions.

summary

Suppose A is 8 × 10 (8 rows and 10 columns) with rank r.

The row space of A is an r-dim subspace of R^{10}.

The col space of A is an r-dim subspace of R^8.

The null space of A (the set of sols to $A\vec{x} = \vec{0}$) is a (10−r)-dim subspace of R^{10}.

how to show that a set of vectors is a subspace

Section 2.5 gave two ways to show that a set of vectors is a subspace:

method 1 Show that it's closed under addition and scalar mult.

method 2 Show that the set is composed of all combinations of a bunch of vectors.

Now you have a third way.

method 3 If the set of vectors is the solution of a homogeneous system of equations then it's a subspace.

example 2

Look at the set of vectors of the form (a,b,c,−a). In example 3 of Section 2.5, I used two methods to show that the set is a subspace. And I found a basis for the subspace. Here's a third way.

The set consists of all vectors (x_1,x_2,x_3,x_4) where $x_4 = -x_1$. So the set is the solution of the homogeneous system

$$x_1 + 0x_2 + 0x_3 + x_4 = 0 \quad \text{(one equation, four unknowns)}$$

So the set is a subspace.

The solution to the system is

$$x_4 = -x_1;\ x_1,x_2,x_3 \text{ free.}$$

There are 3 free variables so the subspace of sols is 3-dim. Here's how to get a basis for the subspace:

Let $x_1=1, x_2=0, x_3=0$ to get sol (1,0,0,−1).
Let $x_1=0, x_2=1, x_3=0$ to get sol (0,1,0,0).
Let $x_1=0, x_2=0, x_3=1$ to get sol (0,0,1,0).

A basis is (1,0,0,−1), (0,1,0,0), (0,0,1,0).

PROBLEMS FOR SECTION 5.2

1. Write the solutions in parametric form and find a basis for the subspace of solutions.

(a) $\left[\begin{array}{ccccc|c} 1 & 5 & 2 & 4 & 0 & 0 \\ 0 & 1 & 3 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(b) $\begin{aligned} x + 2y + z &= 0 \\ x + 3y + 2z &= 0 \end{aligned}$ (c) $\begin{aligned} 2x + y + z &= 0 \\ 4x + y &= 0 \\ -2x + 2y + z &= 0 \end{aligned}$

2. Find a basis for the null space of M.

(a) $M = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 \end{bmatrix}$ (b) $M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ (c) $M = [1\ 1\ 2]$

3. Find an orthogonal basis for the subspace of solutions to

$$\begin{aligned} x + 2y + z - w &= 0 \\ 2x + 5y + 3z - w &= 0 \end{aligned}$$

4. Describe the set of solutions geometrically.

(a) $\left[\begin{array}{ccc|c} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ (b) $\left[\begin{array}{ccc|c} 2 & 3 & 4 & 0 \\ 4 & 6 & 8 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ (c) $\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

(d) $\left[\begin{array}{ccc|c} 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ (e) $\left[\begin{array}{ccc|c} 1 & 0 & -2 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

5. Solve $A\vec{x} = \vec{b}$ where A is the 3×5 zero matrix and

(a) $\vec{b} = \vec{0}$ (b) $b \neq \vec{0}$

6. Suppose A has six cols, the cols are dep but the last five cols are ind. What can you conclude about the number of solutions (and the number of free variables) to $A\vec{x} = \vec{b}$ where (a) $\vec{b} = \vec{0}$ (b) $\vec{b} \neq \vec{0}$

7. Let u = (1,0,2), v = (1,1,4). Find all the vectors orthog to both u and v
(a) using geometry and calculus
(b) using the ideas of this section

8. Let u = (1,1,4,−2). Find as many independent vectors as possible orthogonal to u.

9. Use ideas from this section to show that the set of points is a subspace and find a basis.
(a) The set of points of the form $(x_1, \ldots, x_6)$ where $x_2 = -x_3 + x_5$ and $x_4 = 2x_3 + x_5$
(b) The set of points of the form (a,a,2a,2a,b)

10. Let A and B be fixed 10×7 matrices.

Look at the set of column vectors $\vec{x}$ in R^7 such that $A\vec{x} = B\vec{x}$, i.e., the set of vectors x such that A and B "do the same thing to them". Use ideas of this section to show that this set is a subspace.

11. Solve the system of equations $\left.\begin{matrix} 0 & 2 \\ 0 & 3 \end{matrix}\,\right|\,\begin{matrix} 0 \\ 0 \end{matrix}$ (call the unknowns x and y) and find a basis for the subspace of solutions.

SECTION 5.3 ORTHOGONAL COMPLEMENTS

definition of the orthogonal complement of a subspace

Suppose V is a subspace of R^n.

The set of all vectors orthog to every vector in V is called the orthogonal complement of V and denoted by V^{perp} or $V^{\perp}$.

In other words, u is in $V^{\perp}$ iff u is orthog to every vector in V.

For example, in R^3, a plane through the origin is a 2-dim subspace; its orthogonal complement is the line through the origin perpendicular to the plane (Fig 1). The line consists precisely of those points $\vec{u}$ such that the $\vec{u}$ is perp to every vector in the plane.

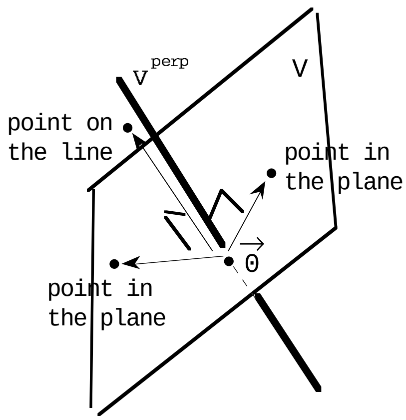


FIG 1

all about the orthog complement

Let V be a subspace of R^n.

(1) $V^{\perp}$ is also a subspace of R^n.

(2) The sum of the dimensions of V and $V^{\perp}$ is n (e.g., if V is a 6-dim sub space of R^{30} then $V^{\perp}$ is a 24-dim subspace of R^{30}).

(3) Not only does $V^{\perp}$ contain everything orthog to all the vectors in V but it can be shown that V in turn contains everything orthog to all the vectors in $V^{\perp}$ so that the spaces V and $V^{\perp}$ are orthog complements of each other.

For example in Fig 1, the line is the orthog complement of the plane and the plane is the orthogonal complement of the line.

(4) Suppose V is spanned by vectors u,v,w,p. To find $V^{\perp}$, let A be the matrix with rows u,v,w,p. Then the set of solutions to $A\vec{x} = \vec{0}$ (the null space of A) is $V^{\perp}$.

(5) The null space and the row space of a matrix are orthog complements.

proof of (4)

Suppose V is spanned by u,v,w,p. To find $V^{\perp}$, find all $\vec{x}$ orthog to the spanning vectors by solving

(*) $$\vec{u}\cdot\vec{x} = 0,\ \vec{v}\cdot\vec{x} = 0,\ \vec{w}\cdot\vec{x} = 0,\ \vec{p}\cdot\vec{x} = 0\ .$$

Once x is orthog to u,v,w,p then x is orthog to every vector in V (see problem #6 in Section 2.5).

So the solution to (*) is the orthog complement of V.

The system in (*) can be written as $A\vec{x} = \vec{0}$ where the rows of A are $\vec{u},\vec{v},\vec{w},\vec{p}$, and $\vec{x}$

is written as a column So to find $V^\perp$, solve $A\vec{x} = \vec{0}$, i.e., find the null space of A.

proof of (5)

This follows from (4) which shows that if a matrix A has rows u,v,w,p then the null space of the matrix is the orthog complement of the space spanned by u,v,w,p.

proof of (1) and (2)

Suppose V is a 6-dim subspace of R^{30}.
Let A be a marix whose rows are a basis for V (so that V is the row space of A).
A is 6×30.

By (4), $V^\perp$ is the null space of A which makes it a subspace, proving (1).
Furthermore,

rank A = 6
number of cols in A = 30
dim of null space = n-r = 24

So dim $V^\perp$ is 24, proving (2).

proof of (3) omitted (subtle)

example 1

Let

$$u = (0,1,0,1,0)$$
$$v = (0,0,1,0,2)$$

Let V be the space spanned by u,v. Find a basis for $V^\perp$

solution Let A be the matrix with rows u,v. Solve the system $A\vec{x} = \vec{0}$:

$$\left[\begin{array}{ccccc|c} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \end{array}\right]$$

It's already in echelon form. The solution is

$$x_3 = -2x_5$$
$$x_2 = -4x_4$$

To find a basis for $V^\perp$, set $x_1 = 1$, $x_4 = 0$, $x_5 = 0$ to get

$$p = (1,0,0,0,0);$$

set $x_1 = 0$, $x_4 = 1$, $x_5 = 0$ to get

$$q = (0,-1,0,1,0);$$

set $x_1 = 0$, $x_4 = 0$, $x_5 = 1$ to get

$$r = (0,0,-2,0,1)$$

A basis for $V^\perp$ is p,q,r.
Note that V and $V^\perp$ are subspaces of R^5 so, by (2), their dimensions should add up to 5. And they do: dim V = 2 (since u and v are ind) and dim $V^\perp$ = 3.

how many perps

Suppose u,v,w are independent vectors in R^{30}.
You can get 27 independent vectors orthogonal to u,v,w, but no more than that.
You can get 28 independent vectors orthogonal to u and v, but no more than that.
You can get 29 independent vectors orthogonal to u, but no more than that.

proof

The number of independent vectors orthog to u,v,w is the same as the dimension of the orthog complement of the subspace spanned by u,v,w. The subspace spanned by u,v,w has dimension 3 (since u,v,w are ind). By (2), the orthog complement has dimension 27. So that's the most independent vectors you can get which are orthog to u,v and w.

Similarly, the number of independent vectors orthog to u is the same as the dimension of the orthog complement of the 1-dim subspace spanned by u. By (2), the orthog complement has dimension 29.

mathematical catechism

question What does it mean to say that U is the orthogonal complement of a subspace W of R^8.

answer It means that U is the set of all vectors in R^8 that are orthogonal to every vector in W.

PROBLEMS FOR SECTION 5.3

1. Find a basis for the orthog complement of the subspace spanned by
   p = (1,1,0,0), q = (0,1,0,1).

2. Find all the vectors perp to
   p = (1,0,0,0,0,1), q = (1,1,0,0,0,0), r = (0,0,1,1,0,0).

3. Let

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 2 & 4 & 1 & 10 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 1 & 2 & 1 & 7 & 0 \end{bmatrix}$$

Find the dimension of and a basis for

(a) the row space of A
(b) the col space of A
(c) the null space of A
(d) the orthogonal complement of the null space of A

4. Let $A = \begin{bmatrix} 2 & 4 & 1 \\ 3 & 6 & 0 \\ 4 & 8 & 0 \\ 5 & 10 & 0 \end{bmatrix}$

(a) Find a basis for the orthogonal complement of the column space.
(b) Continue from part (a) and find an *orthogonal* basis.

5. Do you remember from calculus that the graph in 3-space of 2x + 3y + 4z = 0 is a plane through the origin with normal vector $\vec{n} = 2\vec{i} + 3\vec{j} + 4\vec{k}$. Furthermore, the *only* vectors perp to the plane are $\vec{n}$ and multiples of $\vec{n}$.

Here's the 5-dimensional version of this idea.

Look at the set of points (a "hyperplane") satisfying the equation

$$(*) \qquad 2x_1 + 3x_2 + 5x_3 - 7x_4 + x_5 = 0.$$

A vector is called orthog to the hyperplane if it is orthog to every vector in the hyperplane.

Let

$$\vec{n} = (2,3,5,-7,1)$$

Use the suff from this section to show that $\vec{n}$ is a orthog to the hyperplane and furthermore the only vectors orthog to the hyperplane are $\vec{n}$ and multiples of $\vec{n}$.

6. Let L_1 and L_2 be perpendicular lines through the origin in space. They are subspaces since a line through the origin in R^3 is a subspace. Are they orthogonal complements.

7. Let W be a subspace of R^n. Let w be in W.
Can you decide if w and u are orthogonal if
(a) u is the orthogonal complement of W
(b) u is not in the orthogonal complement of W

8. Suppose (3,1,2) is in the row space of M. Is it possible for (2,1,1) to be in the null space.

8. No. the row space of a matrix and its null space are orthogonal complements. so a vector in the null space has to be orthog to everything in the row space. But (2,1,1) is not orthog to (3,1,2).

SECTION 5.4 SQUARE SYSTEMS

Throughout this section, matrices are square, and all systems of equations have the same number of equations as unknowns.

number of solutions to $M\vec{x} = \vec{b}$

(a) If M is invertible then $M\vec{x} = \vec{b}$ has one solution, namely $\vec{x} = M^{-1}\vec{b}$.

(b) If M is not invertible then $M\vec{x} = \vec{b}$ has either no solutions or infinitely many solutions.

proof of (b)

Suppose M is 3×3 and is not invertible. Then M row ops into something like

$$\begin{matrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix} \quad \text{or} \quad \begin{matrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{matrix} \quad \text{etc.}$$

And $M|b$ row ops to something like

$$\left.\begin{matrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{matrix}\,\right|\,\begin{matrix} 5 \\ 6 \\ 7 \end{matrix}$$

in which case there are no solutions or it row ops to something like

$$\left.\begin{matrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{matrix}\,\right|\,\begin{matrix} 8 \\ 9 \\ 0 \end{matrix}$$

in which case there are infinitely many solutions (because there is a free variable corresponding to the column without a pivot). You can't get exactly one solution because when the system is consistent there will always be at least one free variable.

number of solutions to $M\vec{x} = \vec{0}$

This is a special case of (a) and (b) above.

Remember that a homogeneous system always has at least the trivial solution $\vec{x} = \vec{0}$.

(aa) If M is invertible then $M\vec{x} = \vec{0}$ has *only* the trivial solution $\vec{x} = \vec{0}$.

(bb) If M is not invertible then $M\vec{x} = \vec{0}$ has more (infinitely many more) than just the trivial solution.
I'm going to add this to the invertible rule.

invertible rule.

Let M be $n \times n$.
The following are equivalent; i.e., either all are true or all are false.

(1) M is invertible (nonsingular).
(2) $|M| \neq 0$.
(3) Echelon form of M is I.
(4) Rows of M are independent.
(5) Cols of M are independent.
(6) Rank of M is n.

(7) $Mx = \vec{0}$ has only the trivial solution $x = \vec{0}$.

In other words, the null space of M contains only $\vec{0}$.

In other words, if $x \neq \vec{0}$ then $Mx \neq \vec{0}$

Cramer's rule

Look at $M\vec{x} = \vec{b}$ where M is square.

(1) If $|M| = 0$ then either the system is inconsistent or it has infinitely many solutions (this is just a restatement of (b) from above).

(2) If $|M| \neq 0$ then the system has one solution (this is a restatement of (a)).

Furthermore, in case (2), the solution is $\vec{x} = M^{-1}\vec{b}$ which boils down to

$$x_1 = \frac{\text{replace col 1 of M with } \vec{b} \text{ and then take det}}{\det M}$$

(*)

$$x_2 = \frac{\text{replace col 2 of M with } \vec{b} \text{ and then take det}}{\det M}$$

etc.

proof of the formula in (*)

Here's where the formulas in (*) come from for a 3×3 matrix M.

If $|M| \neq 0$, the sol is $\vec{x} = M^{-1}\vec{b}$. Use the adjoint method for finding M^{-1} to get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{|M|}\begin{bmatrix} \text{cof of } m_{11} & \text{cof of } m_{12} & \text{cof of } m_{13} \\ \text{cof of } m_{21} & \text{cof of } m_{22} & \text{cof of } m_{23} \\ \text{cof of } m_{31} & \text{cof of } m_{32} & \text{cof of } m_{33} \end{bmatrix}^T \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

So

$$x_2 = \frac{1}{|M|}\left(b_1 \cdot \text{cof of } m_{12} + b_2 \cdot \text{cos of } m_{22} + b_3 \cdot \text{cof of } m_{32}\right)$$

$$x_2 = \frac{1}{|M|}\left(-b_1 \begin{vmatrix} m_{21} & m_{23} \\ m_{31} & m_{33} \end{vmatrix} + b_2 \begin{vmatrix} m_{11} & m_{13} \\ m_{31} & m_{33} \end{vmatrix} - b_3 \begin{vmatrix} m_{11} & m_{13} \\ m_{21} & m_{23} \end{vmatrix} \right)$$

$$= \frac{\begin{vmatrix} m_{11} & b_1 & m_{13} \\ m_{21} & b_2 & m_{23} \\ m_{31} & b_3 & m_{33} \end{vmatrix}}{\det M}$$

Similarly for x_1 and x_3.

example 1

The system

$$\begin{aligned} 2x + y + z &= 0 \\ x - y + 5z &= 0 \\ y - z &= 4 \end{aligned}$$

has one solution because

$$\begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & 5 \\ 0 & 1 & -1 \end{vmatrix} = -6 \neq 0$$

The solution is

$$x = \frac{\begin{vmatrix} 0 & 1 & 1 \\ 0 & -1 & 5 \\ 4 & 1 & -1 \end{vmatrix}}{-6} = -4$$

$$y = \frac{\begin{vmatrix} 2 & 0 & 1 \\ 1 & 0 & 5 \\ 0 & 4 & -1 \end{vmatrix}}{-6} = 6$$

$$z = \frac{\begin{vmatrix} 2 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 4 \end{vmatrix}}{-6} = 2$$

PROBLEMS FOR SECTION 5.4

1. Suppose M is 3 × 3 and $M \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

   What important fact(s) can you conclude about M.

2. Try Cramer's rule and if that doesn't help, solve some other way.

(a) $\begin{aligned} 2x - 5y + 4z &= 5 \\ x - 3y + 2z &= 1 \\ y + z &= 7 \end{aligned}$ (b) $\begin{aligned} 2y + 4z &= 10 \\ x + y + 4z &= 4 \\ x + 2z &= 1 \end{aligned}$ (c) $\begin{aligned} 2y + 4z &= 6 \\ x + y + 4z &= 4 \\ x + 2z &= 1 \end{aligned}$

3. Look at the system

$$\begin{aligned} 3x + y &= 3 \\ 2x - y &= 5 \end{aligned}$$

Solve it three times.

(a) Use Cramer's rule.
(b) Use an inverse matrix
(c) Use Gaussian elimination.

4. Use Cramer's rule to solve

$$\begin{aligned} ax + by &= c \\ dx + ey &= f \end{aligned}$$

for x and y. What assumption is necessary for your solution to be valid.

5. Let A be 3 × 3 and let $\vec{b} = \begin{bmatrix} 5 \\ 9 \\ 8 \end{bmatrix}$, $\vec{c} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

   Suppose $A\vec{x} = \vec{b}$ has no solutions.
   What can you conclude about

(a) the number of sols (and the number of free variables) to $A\vec{x} = \vec{0}$
(b) the number of sols (and the number of free variables) to $A\vec{x} = \vec{c}$ where $\vec{c} \neq \vec{0}$

(c) $|A|$
(d) rank A

6. Let $A = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ -2/\sqrt{5} & 1/\sqrt{5} & 0 \\ 1/\sqrt{30} & 2/\sqrt{30} & -5/\sqrt{30} \end{bmatrix}$

Without any agony, solve $A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$.

CHAPTER 6 LINEAR OPERATORS

SECTION 6.1 MATRIX OPERATORS

linearity

Let T be a function (i.e., operator, transformation) which maps input vectors to output vectors. T is called *linear* if

(1) $$\boxed{T(u + v) = T(u) + T(v) \quad \text{and} \quad T(au) = aT(u)}$$

for all vectors u,v and all scalars a.

Linear operators are the mathematical models for physical input-output devices that satisfy the *superposition principle* where the response to a sum of inputs is the sum of the separate responses, and tripling an input for instance will triple the response.

matrix operators

Let

$$M = \begin{bmatrix} 3 & 0 & 0 \\ 4 & -5 & 6 \\ 2 & 0 & 1 \end{bmatrix}$$

Then

$$M\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 6 \\ 23 \\ -1 \end{bmatrix}$$

and in general

$$M\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x \\ 4x-5y+6z \\ 2x-y \end{bmatrix}$$

We say that M maps input (2,3,5) to output (6,23,−1) and in general M maps (x,y,z) to (3x, 4x−5y+6z,2x−y). Inputs are also called *pre-images*, and outputs are called *images*.

The matrix M is an operator on R^3.

In general, any n × n matrix M is an operator on R^n and is linear since

$$M(\vec{u} + \vec{v}) = M\vec{u} + M\vec{v} \quad \text{and} \quad M(k\vec{u}) = k(M\vec{u})$$

for all col vectors $\vec{u},\vec{v}$ and all scalars k (by rules of matrix algebra). Not only are matrix operators linear but it can be shown that they are the only linear operators on R^n.

Non-square matrices are also linear operators. If

$$M = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix}$$

then M maps from R^3 to R^2. For example

$$M\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 22 \end{bmatrix}$$

so M sends (3,0,1) to (10,22).

As an example of a *non*-linear operator on R^3, let

$$T(x,y,z) = (2x,\ y+z,\ 7).$$

One way to show that T is not linear is to find a counterexample to one of the properties in (1):For instance

$T(i+j) = T(1,1,0) = (2,1,7)$
$T(i) + T(j) = (2,0,7) + (0,1,7) = (2,1,14)$.
$T(i+j) \neq T(i) + T(j)$.
So T does not satisfy the first property in (1) so T is not linear.

There is no matrix M (whose entries are specific numbers, not variables) such that

$$M \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ y+z \\ 7 \end{bmatrix}$$

which is another way to recognize that T is not linear.

example 1

If point (x,y) with polar coordinates r, α is rotated ccl by angle θ (Fig 1) then the new point has polar coordinates r, $\theta + \alpha$. So

$$\text{new } x = r\cos(\alpha+\theta) = \underbrace{r\cos\alpha}_{\text{old } x}\cos\theta - \underbrace{r\sin\alpha}_{\text{old } y}\sin\theta$$

Similarly it turns out that

$$\text{new } y = (\text{old } x)\sin\theta + (\text{old } y)\cos\theta$$

So the rotation sends (x,y) to $(x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta)$.
In matrix notation,

$$(2) \qquad \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{\substack{\text{rotation operator} \\ \text{DO NOT MEMORIZE}}} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\text{in}} = \underbrace{\begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}}_{\text{out}}$$

So rotation in 2-space by angle θ is a linear operator, in particular it is multiplication by the indicated matrix in (2).

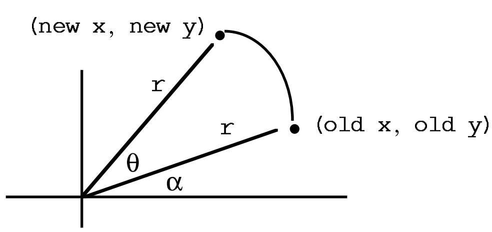


FIG 1

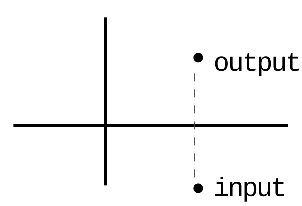


FIG 2

example 2

Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Then A maps point (x,y) to point (x,−y) so it reflects points in the x-axis (Fig 2).

inverse of a matrix operator

Suppose A is an invertible matrix. If A is thought of as a transformation on R^n then A^{-1} is the inverse transformation; i.e., it maps backwards.

For example, let

$$M = \begin{bmatrix} \cos 27^\circ & -\sin 27^\circ \\ \sin 27^\circ & \cos 27^\circ \end{bmatrix}$$

By (2), this rotates points counterclockwise by 27°. Using the standard method for inverting a 2×2 matrix (Section 1.5), we have

(3) $$M^{-1} = \begin{bmatrix} \cos 27^\circ & \sin 27^\circ \\ -\sin 27^\circ & \cos 27^\circ \end{bmatrix}$$

Let's check that this matrix does map backwards, i.e., rotates *clockwise* by 27°, i.e., rotates *counterclockwise* by -27°: By (2), the matrix that does rotate counterclockwise by -27° is

(4) $$\begin{bmatrix} \cos(-27^\circ) & -\sin(-27^\circ) \\ \sin(-27^\circ) & \cos(-27^\circ) \end{bmatrix}$$

Use a little trig to see that (3) and (4) are the same. So the inverse in (3) does map backwards.

product of matrix operators

If matrices A and B are viewed as transformations then the product AB is a composition of the transformations. In particular B goes first and A goes second because when you find $AB\vec{x}$, B multiplies $\vec{x}$ first.

For example, if

$$C = \begin{bmatrix} \cos 60^\circ & -\sin 60^\circ \\ \sin 60^\circ & \cos 60^\circ \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

then

$$DC = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}$$

The matrix C rotates points ccl by 60° and D reflects points in the x-axis so DC first rotates points by 60° and then reflects them in the x-axis (Fig 3).

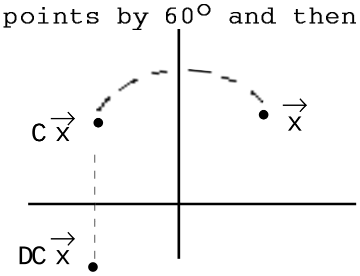


FIG 3

finding a matrix operator given what it does to a basis

Here are two examples to illustrate the idea.

(a) Let $\vec{p} = (2,3,4)$, $\vec{q} = (5,6,7)$, $\vec{r} = (5,5,5)$.

Find the matrix M that maps $\vec{i}$, $\vec{j}$, $\vec{k}$ to $\vec{p}, \vec{q}, \vec{r}$ respectively. In other words, find M so that Mi = p, Mj = q, Mk = w.

solution Put p,q,r down the cols, i.e., let

$$M = [p\ q\ r] = \begin{bmatrix} 2 & 5 & 5 \\ 3 & 6 & 5 \\ 4 & 7 & 5 \end{bmatrix}$$

This works because by observation (4) about matrix multiplication in Section 1.1, $M\vec{i}$ is col 1 of M, $M\vec{j}$ is col 2 of M etc.

footnote

From one point of view, M is an operator that sends $\vec{i}$, $\vec{j}$, $\vec{k}$ to $\vec{p},\vec{q},\vec{r}$ resp. From another point of view, M is the basis changing matrix (that I usually call P), that converts the coords of a point w.r.t. basis $\vec{p},\vec{q},\vec{r}$ to the coordinates of that point w.r.t. basis $\vec{i}$, $\vec{j}$, $\vec{k}$.

(b) Let u,v,w be independent vectors in R^3. Let p,q,r be in R^3 (not necessarily independent). Find the matrix M that maps u,v,w to p,q,r respectively. In other words, find M so that Mu = p, Mv = q, Mw = r.

footnote If you allow u,v,w to be dependent then there might not be such an M.
For instance if u = (1,1,1), v = (2,2,2) (note that v=2u), p = (5,6,7), q = (3,2,5) then you can't find a matrix M such that Mu = p and Mv = q. Once Mu = p then Mv = M(2u) = 2Mu = 2p, not q.
Or it might turn out that there is more than one such M.
In any event, I'm only asking the question for *independent* u,v,w.

solution First find the matrix A that maps i,j,k to u,v,w. By part (a),

A = [u v w] (i.e., the cols are A are u,v,w).

And find the matrix B that maps i,j,k to p,q,r. By part (a),

B = [p q r]

Then A is invertible because its cols are ind and A^{-1} maps u,v,w to i,j,k.

Put this all together:

A^{-1} maps u,v,w to i,j,k respectively.
B maps i,j,k to p,q,r respectively.

So BA^{-1} maps u,v,w to p,q,r respectively.

So the answer is $M = BA^{-1}$.

range of a matrix operator

The range of matrix A is the set of all outputs.

If say $A = \begin{bmatrix} 2 & 3 & 4 \\ 7 & 8 & -9 \end{bmatrix}$ then vector $\vec{b}$ is in the range if there is a vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ such that

$$\begin{bmatrix} 2 & 3 & 4 \\ 7 & 8 & -9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \vec{b}$$

This can be written as

$$x \begin{bmatrix} 2 \\ 7 \end{bmatrix} + y \begin{bmatrix} 3 \\ 8 \end{bmatrix} + z \begin{bmatrix} 4 \\ -9 \end{bmatrix} = \vec{b}$$

So

$\vec{b}$ is in the range of matrix A

iff the equation $A\vec{x} = \vec{b}$ has a solution (at least one solution)

iff $\vec{b}$ is a combination of the cols of A

iff $\vec{b}$ is in the col space of M.

In general:

The range of M is a subspace; in particular it's the column space of M.

The range of M is also the set of $\vec{b}$'s that make the equation $M\vec{x} = \vec{b}$ consistent.

If M is say 7×9 with rank r then M maps from R^9 to R^7 and the range is an r-dim subspace of R^7.

rank list

Given a matrix M, the following are all equal and are called the rank of M:

(1) maximal number of ind cols (i.e., dim of the col space of M)
(2) maximal number of ind rows (i.e., dim of the row space of M)
(3) number of cols with pivots in the echelon form of M
(4) number of nonzero rows in the echelon form of M
(5) largest order of all the nonzero subdeterminants of M
(6) dimension of the range of transformation M

warning

Don't confuse the rank with the range. The rank of a matrix is a number. The range of a matrix is a subspace (the col space). The rank is the *dimension* of the range.

null space of a matrix operator

The null space of M is the set (actually a subspace) of solutions to $M\vec{x} = \vec{0}$ (§5.2).

So the null space of the operator M is the set of vectors mapping to $\vec{0}$.

If M has n cols and rank r then the null space of M is an (n−r)-dim subspace of R^n

(§5.2) and its range has dimension r so

dim of null space of M + dim of range of M = n

example 3

Let

$$M = \begin{bmatrix} \cos 25^\circ & -\sin 25^\circ \\ \sin 25^\circ & \cos 25^\circ \end{bmatrix}$$

Then M rotates points in R^2 ccl by 25°.

The null space contains only $\vec{0}$ since only the origin rotates into the origin.
The range contains every vector in R^2 (i.e., the range *is* R^2) since every point is rotated into (by the point 25° before it).

summary

Suppose M has 6 rows and 9 cols with rank 4 (4 echelon cols have pivots).

Then M maps from R^9 to R^6 (M can multiply a col vector that has 9 entries and produces a col vector that has 6 entries).

The range (same as the col space) of M is a 4-dim subspace of R^6.

The null space of M is a 5-dim subspace of R^9 (the system of equations Mx=0 has 9 variables and 5 of them are free because there are 5 echelon cols without pivots).

The row space of M is a 4-dim subspace of R^9.

In general (but I wouldn't memorize most of this, I would figure it out each time):

Let M be $m \times n$ (i.e., m rows, n cols) with rank r.

As an operator, M maps from R^n to R^m.

The col space of M which is the same as the range of M is an r-dim subspace of R^m.

The null space of M is an (n−r)-dim subspace of R^n.

The row space of M is an r-dim subspace of R^n (same dim as the col space).

warning

1. If M is 5×5 and rank M = 3 then the range of M is a 3-dim subspace of R^5. The range is *not* R^3 itself. The outputs of the matrix are *five*-tuples, not *three*-tuples.
2. If M is 5×5 and rank M = 5 then the range is a 5-dim subspace of R^5. Since the only 5-dim subspace of R^5 is R^5 itself, the range in this case *is* R^5.

invertible rule

Let M be $n \times n$.

The following are equivalent; i.e., either all are true or all are false.

(1) M is invertible (nonsingular).
(2) $|M| \neq 0$.
(3) Echelon form of M is I.
(4) Rows of M are independent.
(5) Cols of M are independent.
(6) Rank of M is n.
(7) $Mx = \vec{0}$ has only the trivial solution $x = \vec{0}$.

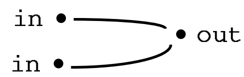
In other words, the null space of M contains only $\vec{0}$.

In other words, if $x \neq \vec{0}$ then $Mx \neq \vec{0}$.

In other words, the operator M maps $\vec{0}$, *and nothing else*, to $\vec{0}$.

In other words, M can't map a nonzero vector to $\vec{0}$.

(8) The operator M is one-to-one, i.e., M *doesn't* send two inputs to the same output. (Fig 4).



impossible for an invertible matrix
FIG 4

(9) The range of M is R^n.

Item (8) can be added to the list because M is invertible iff operator M has an inverse operator which happens iff M is a one-to-one operator.

Item (9) can be added because it is equivalent to (6): The range is all of R^n iff the dimension of the range is n; but the dim of the range or M *is* the rank of M.

rank of a product

The rank of a product of two matrices is ≤ the rank of either factor. Furthermore if one of the factors is invertible then the product rank equals the rank of the other factor; i.e., multiplying by an invertible matrix preserves rank.

In other words:

(1) rank AB ≤ rank A.
(2) rank AB ≤ rank B.
(3) If B is invertible then rank AB = rank A.
(4) If A is invertible then rank AB = rank B.

warning

Multiplying by a non-invertible matrix might also preserve rank but you can't count on it. In other words, if B is not invertible then rank AB *might* still equal rank A. But a non-invertible B can't preserve the rank of *every* A. It can be shown that only invertibles can do that.

proof of (1)

Here's the proof in the case that A and B are both $n \times n$. A similar argument works for non-squares.

Fig 5 shows the operator A.
Fig 6 shows the operator AB.

Each picture shows all of R^n going in initially.
I want to compare the dimensions of the two sets of outputs.

In Fig 5, all of R^n goes directly into the A machine

In Fig 6, the A machine does not necessarily get all of R^n as input; it gets only the outputs from the B machine. Because the flow into A might be diminished by having to pass through B first, the set of outputs in Fig 6 is a subset of the set of outputs of Fig 5.

In other words, range AB $\subseteq$ range A.
Since rank is the dimension of the range, rank AB ≤ rank A.

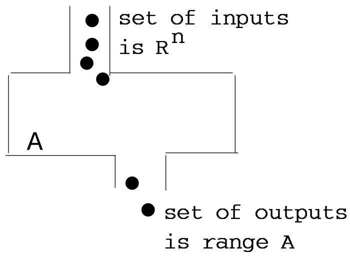


FIG 5 Operator A

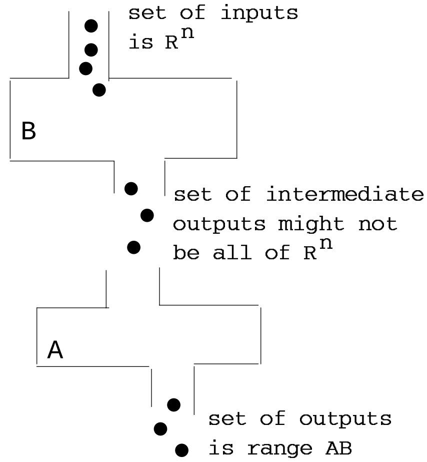


FIG 6 Operator AB

proof of (3)

Continue with the pictures in Figs 5,6. Suppose B is invertible. Then the set of outputs from the B machine is *all* of R^n (this is (11) on the invertible list). So in each picture, all of R^n goes into the A machine, so the two sets of outputs are the same. So if B is invertible then range AB = range A and rank AB = rank A.

proof of (2)

rank AB = rank$(AB)^T$	transposes have the same rank
= rank(B^TA^T)	T rule
$\le$ rank B^T	by rank of product rule (1) which was just proved
= rank B	transposes have the same rank

proof of (4)

If A is invertible then so is A^T and

rank AB = rank(B^TA^T)	as above
= rank B^T	by rank of product rule (3) since A^T is invertible
= rank B	transposes have the same rank

PROBLEMS FOR SECTION 6.1

1. Suppose a transformation sends (x_1,x_2,x_3,x_4) to $(x_4,\ x_4,\ 2x_1-3x_2,\ 5x_1+6x_3+7x_4)$. Find the matrix of the transformation.

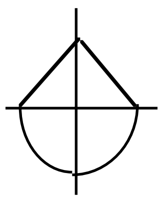
2. Let $M = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$.
Find the image of (2,4) and the pre-image of (0,1).

3. Let

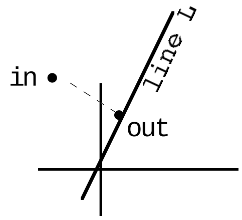
$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},\quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},\quad C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

(a) What does each operator do geometrically and in particular what do they do to the figure in the diagram.
(b) Find the inverse of each matrix (if an inverse exists), not by computing an inverse but by thinking about mappings.

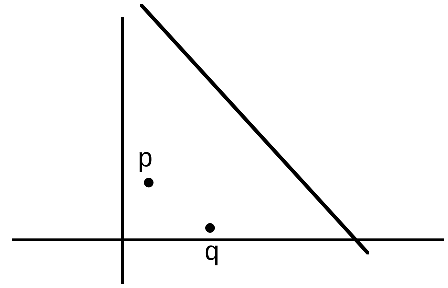
(c) Find B^{100} by thinking geometrically about the mapping B.
(d) Find C^2 by thinking geometrically about the mapping B.
(e) Find the range of each matrix operator.



Problem 3(a)



Problem 4



Problem 5

4. Let L be the line in 2-space have equation $y = 7x$ (see the diagram). Find the matrix M which projects points onto L.

5. The purpose of this problem is to show that reflection in a line *not* through the origin is *not* a linear operator.
  Let T reflect in the line *not* through the origin in the diagram.
 (a) The diagram also shows points p and q. Plot points
  $p+q$, $T(p+q)$, $T(p)$, $T(q)$ and $T(p) + T(q)$.
 (b) Use part (a) to show that T is not linear.

6. Find a 2×2 matrix which rotates points in R^2 by 60° and then reflects them in the y-axis.

7. Suppose M is 4×4 so that M maps from R^4 to R^4. Let $\vec{v} = (1,2,9,4)$. Is it possible for M to map exactly 13 vectors to $\vec{v}$?

8.(a) Find the matrix A that maps $\vec{i}$ to $(1,4,2)$, $\vec{j}$ to $(5,3,\pi)$, $\vec{k}$ to $(-2,7,8)$.

 (b) Let $u = (1,2,3)$
  $v = (0,1,1)$
  $w = (1,1,1)$
  $q = (5,3,2)$
  $r = (1,2,1)$
  $s = (3,3,0)$

I checked that u,v,w are ind.
Find the matrix that maps u,v,w to q,r,s respectively.
(It's OK to leave the answer in a form such as $PQRS^{-1}M^T$ as long as you say what P,Q,R,S,M are.)

9. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(a) Decide what each matrix operator does geometrically and then use this to find the range, null space and rank of each matrix.
(b) Find the null space of A algebraically to see if it agrees with your answer in part (a).

10. Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Describe the range of each geometrically and sketch. Find a basis for each range.

11. Let

$$M = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \text{ and } u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

(a) Is u in the null space of M?
(b) Is u in the range of M?

12. Let $M = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 2 \\ 2 & 0 & 4 \\ 1 & 1 & 5 \end{bmatrix}$.

Find a basis for (a) the null space of M (b) the range of M

13. Suppose M is 10×10.
Describe the range, null space, determinant if the rank of M is (a) 10 (b) 7

14. Suppose M is 5×9. Describe the range and null space of M if the rank of M is (a) 5 (b) 3

15. Let M be 6×6. True or False.
(a) If rank M = 4 then range of M is R^4.
(b) If rank M = 6 then range of M is R^6.

16. Let $b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$. Let A be 4×4.

(a) Is it possible for A to produce only the output b and nothing else.
(b) Suppose b is an output of A. Can you tell *how many* inputs produce b, i.e., how many pre-images might b have. (Is it just one, or maybe two or maybe ten? What are the possibilities?)
(c) Suppose A can not only produce output b, but every other vector in R^4 as well. How many possible inputs produce b?

17. Can row operations change
(a) the null space of a matrix operator
(b) the range of a matrix operator

18. (a) If possible, find a matrix whose outputs are all the vectors of the form (a,a,0,b). In other words, find a matrix that can output vectors like (2,2,0,7), $(\pi,\pi,0,\sqrt{2})$ etc. and only vectors like this.
(b) If possible, find a matrix whose outputs are all the vectors of the form (a,a,2,b)

19. Suppose rank A = 4 and rank B = 7 (and the product AB exists).
(a) Draw the best conclusion you can about rank(AB).
(b) Draw the best conclusion you can about rank(AB) if B is invertible.
(c) Suppose A were invertible. Then rank(AB) $\le$ 4 from part (a), but rank(AB)=7 since multiplying B by an invertible matrix preserves rank. Impossible. What went wrong.

20. Suppose A is 3×4, rank A = 3 and AB = I. Find the rank of B.

21. Suppose A is 7×9 (7 rows, 9 cols) with rank 4.

(a) The range of A is a ?–dim subspace of $R^?$.

(b) Suppose $\vec{x}_1, \ldots, \vec{x}_6$ are independent vectors in R^9.

Use part (a) to show that the vectors $A\vec{x}_1, \ldots, A\vec{x}_6$ are dependent.

22. Find the matrix M that maps u = (4,0) to v = (1,3), and p = (6,1) to q = (2,5). After you find M, check numerically that you really do have Mu = v and Mp = q.

SECTION 6.2 BASIS CHANGING

the matrix of a linear transformation w.r.t. a basis

Suppose a linear transformation T on R^2 sends point $\vec{p}$ to point $\vec{q}$.
If say p = (3,−1) and q = (2,17) (coords w.r.t. basis i,j) and A is the matrix of the transformation then

$$A \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 17 \end{bmatrix}$$

In particular, A is the matrix of the transformation *w.r.t. basis i, j.*

Suppose you use a new coord system determined by basis u,v. Suppose $\vec{p}$ has coords 8,12 and q has coords 5,3 w.r.t. basis u,v; i.e., p = 8u + 12v, q = 5u + 3v. If B is the matrix of the transformation *w.r.t. basis u,v* then

$$B \begin{bmatrix} 8 \\ 12 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

basis changing rule

(1)

Suppose a transformation T mapping from R^3 to R^3 has matrix A w.r.t. basis i,j,k. Let

$$\vec{u} = (u_1, u_2, u_3), \quad \vec{v} = (v_1, v_2, v_3), \quad \vec{w} = (w_1, w_2, w_3)$$

be a new basis for R^3. (The same idea works in R^n.)

$$\text{Let } P = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} \quad \text{(the usual basis changing matrix)}$$

Then the matrix of the transformation w.r.t. basis u,v,w is $P^{-1}AP$.

proof of (1)

Suppose T sends $\vec{x}$ to $\vec{y}$.

Suppose $\vec{x}$ has standard coordinates x_1, x_2, x_3 w.r.t. basis i,j,k and has coords X_1, X_2, X_3 w.r.t. basis u,v,w.

Suppose $\vec{y}$ has standard coords y_1, y_2, y_3 and new coords Y_1, Y_2, Y_3. Then

$$A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \qquad \text{because A is the matrix of the trans w.r.t. i,j,k}$$

$$AP \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = P \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} \quad \text{because} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = P \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = P \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} \quad \text{see (3) in §2.4)}$$

$$P^{-1}AP \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} \qquad \text{left multiply on each side by } P^{-1}$$

So $P^{-1}AP$ is the matrix of the transformation w.rt. basis u,v,w.

example 1

Suppose the matrix of a linear transformation T (w.r.t. the standard basis i,j,k) is

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Let

u = (1,0,0)
v = (1,1,0)
w = (0,0,2)

Find the matrix B that represents the transformation w.r.t. basis u,v,w.

solution Let

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Then

$$P^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$$

and the new matrix of T is

$$B = P^{-1}AP = \begin{bmatrix} 2 & 2 & -2 \\ 0 & 0 & 2 \\ 0 & 1/2 & 0 \end{bmatrix}$$

example 1 continued

Use the matrix B to find T(u + 2w)

solution The vector u + 2w has coordinates 1,0,2 w.r.t. basis u,v,w.

$$B \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix}$$ so T(u+2w) = −2u + 4v.

example 1 continued again

Use the matrix A to find T(u + 2w).

solution

u+2w = (1,0,0) + (0,0,4) = (1,0,4).

$$A \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}$$ so T(u+2w)= T(i+k) = 2i + 4j.

Here's a check that the two answers for T(u+2w) agree.

First answer = −2u + 4v = −2(1,0,0) + 4(1,1,0) = (2,4,0) = second answer.

example 1 continued some more

Find the old "formula" for T and the new "formula".

solution

$$A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ z \\ y \end{bmatrix}$$ so T sends (x,y,z) to (2x,z,y).

So the old formula is T(x,y,z) = (2x,z,y).
On the other hand,

$$B \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2X+2Y-2Z \\ 2Z \\ \frac{1}{2}Y \end{bmatrix}$$

so the new formula for T is $T((X,Y,Z)) = ((2X + 2Y - 2Z,\ 2Z,\ \frac{1}{2}Y))$.

similar matrices

If $B = Q^{-1}AQ$ for some Q (or equivalently $A = R^{-1}BR$ for some R) then B and A are called *similar* matrices.

mathematical catechism

Question 1 What does it mean to say that C is the matrix of a transformation w.r.t. basis u,v,w.

Answer It means that

$$C \begin{bmatrix} \text{coords} \\ \text{of an} \\ \text{input} \\ \text{w.r.t.} \\ \text{u,v,w} \end{bmatrix} = \begin{bmatrix} \text{coords} \\ \text{of the} \\ \text{output} \\ \text{w.r.t.} \\ \text{u,v,w} \end{bmatrix}$$

question 2 What does it mean to say that A and B are similar matrices.

answer 1 It means there exists a matrix Q such that $A = Q^{-1}BQ$

answer 2 It means there exists a matrix R such that $B = R^{-1}AR$.

PROBLEM FOR SECTION 6.2

1. A transformation has matrix $M = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$ (meaning w.r.t. basis i,j).
Find its new matrix w.r.t. basis u = (3,0), v = (−1,2).
Find the old and new transformation formulas, i.e., what does the transformation do to (x,y) and to ((X,Y)).

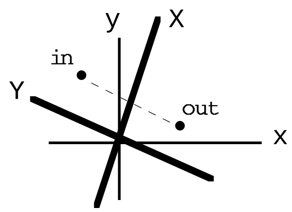
2. A linear transformation T has matrix $M = \begin{bmatrix} 3 & 1 \\ 4 & -1 \end{bmatrix}$ w.r.t. basis u = (2,0), v = (2,1).
(a) Find the matrix (call it A) for T w.r.t. basis i,j.
(b) Find T(2u + v) using the matrix M.
(c) Find T(2 u + v) again, using the matrix A from part (a). And then check that this answer agrees with the answer to (b).

3. Let $M = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{bmatrix}$.
Suppose M is the matrix of a transformation T w.r.t. basis u = (1,4,6), v = (−8,1,−2), w = (2,2,2).
(a) Find T(u).
(b) Find the matrix of T w.r.t. i,j,k.

4. The problem is to find the matrix which reflects points in line y = 3x (see the diagram).
One way to do it is to switch to a new orthog coord system in which the line is one of the axes, solve the problem in the new coord system (it's easy to find a matrix which reflects in a coordinate axis), and then switch the answer back to the old system.
In particular, let T be the transformation which reflects points in line y = 3x and go through the following steps.

step 1 Find a pair of basis vectors for a new orthogonal system in which the line $y=3x$ is say the X-axis.
step 2 Find the matrix (call it B) for the transformation w.r.t. the new basis.
step 3 Convert your answer from step 2 to get the matrix (call it A) of the transformation w.r.t. the old basis.

As a check, use the matrix A to find the reflection of point (−5,5) and see if it looks right.



Problem 4

5. Suppose T operates on R^2 and sends (x,y) to $(x,-y)$. Let $u = (1,4)$, $v = (-5,2)$.

(a) Find the matrix, (call it A) of T w.r.t. basis i,j.
(b) Find the matrix (call it B) of T w.r.t. basis u,,v.
(c) In the new coord system with basis u,v, T sends ((X,Y)) to ((what, what)).

6. Let operator T send (x,y) to (y,x).
(a) What does T do geometrically.
(b) Find the matrix of T w.r.t. the basis $u = (2,1)$, $v = (1,0)$.
(c) Use part (b) to find $T(2u-v)$. Then plot $2u-v$ and $T(2u-v)$ in the u,v coord system and see if the points *look* right.

7. (a) Let

$$\begin{aligned}
p &= p_1 i + p_2 j + p_3 k\\
q &= q_1 i + q_2 j + q_3 k\\
r &= r_1 i + r_2 j + r_3 k,\\
u &= u_1 i + u_2 j + u_3 k\\
v &= v_1 i + v_2 j + v_3 k\\
w &= w_1 i + w_2 j + w_3 k
\end{aligned}$$

A transformation has matrix M_{pqr} w.r.t. basis p,q,r and has matrix M_{uvw} w.r.t. basis u,v,w.

Express M_{uvw} in terms of M_{pqr}

(b) Show that M_{uvw} and M_{pqr} are similar matrices.

8. Show that similar matrices have the same rank.

9. Show that if A is similar to B then A^T is similar to B^T.

REVIEW PROBLEMS FOR CHAPTERS 5 AND 6

1. Look at the system of equations

$$\begin{aligned} a - b - c &= 1 \\ a + b + c &= 0 \\ a - b + c &= 0 \\ a + b - c &= 0 \end{aligned}$$

(a) Use Gaussian elimination to see that it has no solution for a,b,c.
(b) Fill in the blanks.
Since the system has no solution, the vector ____ is not in the space spanned by the vectors __________________ .

(c) Find the least squares solution.
(d) The least squares solution turned out to be $a = \frac{1}{4}$, $b = -\frac{1}{4}$, $c = -\frac{1}{4}$. These values don't satisfy the original system of equations. So what good are they?

2. Suppose A and B are similar and B is invertible. Show that A is also invertible.

3. Let M be the 4×4 zero matrix. Find det M, rank M, range M and null space of M.

4. Find a matrix A if possible (it can be any size you like) so that
(a) $A\vec{x} = \vec{b}$ has no solutions or one solution depending on $\vec{b}$
(b) $A\vec{x} = \vec{b}$ has infinitely many solutions no matter what $\vec{b}$ is
(c) $A\vec{x} = \vec{b}$ has no solutions or infinitely many sols depending on $\vec{b}$
(d) $A\vec{x} = \vec{b}$ has one solution no matter what $\vec{b}$ is
(e) $A\vec{x} = \vec{b}$ has no solutions no matter what $\vec{b}$ is

5. Suppose rank A = 6. What can you you conclude about rank AB if
(a) B is invertible (b) B is not invertible

6. The echelon form of M is

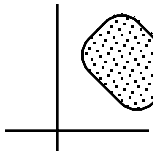
$$\begin{matrix} 1 & 2 & 3 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

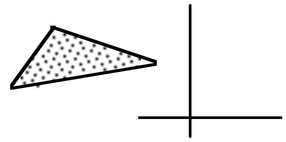
Do whichever of the following are possible with this information (remember that the questions are about the *original* matrix M).
(a) Find rank M.
(b) Find the dimension and a basis for the range of M.
(c) Find the dimension and a basis for the null space of M.
(d) Let u = (1,2,3,4,5,6). Is u in the null space of M.
(e) Find the 4×4 subdeterminant in the northwest corner of M.
(f) Find the 3×3 subdeterminant in the southwest corner of M.
(g) Solve $M\vec{x} = \vec{b}$ where $b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$.

7. Suppose A represents a linear transformation mapping R^n to R^n w.r.t. the standard basis and B represents the same transformation w.r.t. a new basis.
Show that A and B have the same determinant.

8. Suppose M is a matrix operator on R^2 and $|M| = 0$.
A student found the image of the region in the lefthand diagram and concluded that it was the region in the righthand diagram.

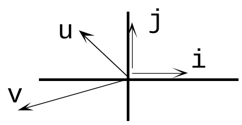
Why is that impossible.





Problem 8

9. Let A be the matrix of a transformation w.r.t. basis u,v (sketched in the diagram) and let B be the matrix of the same transformation w.r.t. basis i,j.



Problem 9

(a) Suppose $A\begin{bmatrix}3\\2\end{bmatrix} = \begin{bmatrix}1\\-1\end{bmatrix}$

Sketch the corresponding input and output vectors in the diagram.

(b) Suppose $B\begin{bmatrix}3\\2\end{bmatrix} = \begin{bmatrix}0\\3\end{bmatrix}$

Sketch the corresponding input and output vectors in the diagram.

(c) What's the algebraic connection between the matrices A and B.

10. Let u = (1,3,3,2), v = (2,6,9,5), w = (−1,−3,3,0). Find a maximally large collection of independent vectors orthogonal to u,v,w.

CHAPTER 7 COMPLEX VECTORS AND MATRICES

SECTION 7.1 COMPLEX NUMBERS

form of a complex number

Expressions of the form a + bi where a and b are real numbers and $i^2 = -1$ are called complex numbers. The real part is a and the imaginary part is b (the imaginary part is b, *not* bi). If the real part is 0 then the number is pure imaginary (e.g., 3i, i, $-\pi$i). If the imaginary part is 0 then the number is real. The complex numbers include the reals as a special case.

conjugation

If z = a + bi then $\overline{z}$ = a − bi (the conjugate of z)
For example,

if z = 6 − 3i then $\overline{z}$ = 6 + 3i;
if z = i then $\overline{z}$ = −i;
if z = 4 then $\overline{z}$ = 4

addition, multiplication, division

If

$$z = 2 + 3i \text{ and } w = 4 - 5i$$

then

$$z + w = 6 - 2i$$

$$zw = (2+3i)(4-5i) = 8 - 15i^2 + 12i - 10i = 23 + 2i$$

$$\frac{z}{w} = \frac{2+3i}{4-5i}$$

$$= \frac{2+3i}{4-5i} \cdot \frac{4+5i}{4+5i} \quad \text{multiply up and down by the conjugate of the denom}$$

$$= \frac{-7 + 22i}{41}$$

$$= -\frac{7}{41} + \frac{22}{41} i$$

the reciprocal of i

$$\frac{1}{i} = \frac{1}{i}\,\frac{i}{i} = \frac{i}{-1} = -i$$

magnitude (modulus)

If z = x + iy then $|z| = \sqrt{x^2 + y^2}$.
For example,

if z = 2 + 3i then $|z| = \sqrt{13}$;
if z = 6i then |z| = 6;
if z = −4 then |z| = 4

If z is real then the mag of z is its absolute value; i.e., the mag of a complex number generalizes the idea of absolute value of a real number.

If z = x + iy is pictured as the point (x,y) then |z| is the distance from z to the origin, i.e., it's the polar coord r.

rules of complex algebra

(1) $\overline{z + w} = \overline{z} + \overline{w}$

(2) $\overline{zw} = \overline{z}\,\overline{w}$

(3) $z\overline{z} = |z|^2$

In particular, if z = a + bi then $z\overline{z} = |z|^2 = a^2 + b^2$.

(4) $\overline{\overline{z}} = z$

(5) $|zw| = |z|\,|w|$

(6) $\left|\dfrac{z}{w}\right| = \dfrac{|z|}{|w|}$

proof of (5)

Let $z = a + bi$, $w = c + di$. Then

$$zw = ac - bd + (bc + ad)i$$

and

$$|zw| = \sqrt{(ac-bd)^2 + (bc+ad)^2} \qquad \text{definition of mag}$$

$$= \sqrt{a^2c^2 - 2abcd + b^2d^2 + b^2c^2 + 2abcd + a^2d^2}$$

On the other hand

$$|z|\,|w| = \sqrt{a^2 + b^2}\ \sqrt{c^2 + d^2}$$

$$= \sqrt{a^2c^2 + b^2d^2 + b^2c^2 + a^2d^2}$$

Look closely and you'll see that $|zw| = |z|\,|w|$

proof of (1)

Let $z = a + bi$, $w = c + di$. Then

$$z + w = a+c + (b+d)i$$

$$\overline{z + w} = a+c - (b+d)i$$

$$\overline{z} + \overline{w} = a-bi + c-di = a+c - (b+d)i$$

So $\overline{z + w} = \overline{z} + \overline{w}$.

PROBLEMS FOR SECTION 7.1

1. If $z = 5 - 3i$ and $w = 4 - 6i$ find $|z|$, $\overline{z}$, zw, $1/z$, z^2.

2. Suppose $z = \overline{z}$. What can you conclude about z.

3. Suppose $z = -\overline{z}$. What can you conclude about z.

4. If $A = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$ (where i is the imaginary number) find A^{243}, A^{244}, A^{245}.

5. Find the magnitude.
(a) $2-7i$ (b) $-i$ (c) 3 (d) $2i$ (e) -7 (f) i (g) $4 + 5i$ (h) $\cos\theta + i\sin\theta$

6. Let $z = 2 - i$ and $w = -3 + 4i$.

(a) Find zw, $\overline{zw}$, $\overline{z}$, $\overline{w}$ and check that $\overline{zw}$ really does equal $\overline{z}\,\overline{w}$.
(b) Find zw, $|zw|$, $|z|$, $|w|$ and check that $|zw|$ really does equal $|z||w|$.

7. Let $z = \pi + \sqrt{2}\,i$. If z is multiplied by $\sqrt{17} + i$, find the magnitude of the result (with as little effort as possible).

SECTION 7.2 HERMITIAN AND SKEW HERMITIAN MATRICES

the conjugate transpose

The conjugate transpose of A, denoted by A^*, is found by conjugating each entry and transposing; i.e., the ij-th entry in A^* is the conjugate of the ji-th entry in A.

If

$$A = \begin{bmatrix} 2 & i \\ 3+i & -6i \end{bmatrix}$$

then

$$A^* = \begin{bmatrix} 2 & 3-i \\ -i & 6i \end{bmatrix}$$

If all the entries in A are real then A^* is just A^T.
A^* is also denoted by A^H or $\overline{A}^T$ and is sometimes called the *Hermitian transpose*.

properties of the conjugate transpose

(1) $(A + B)^* = A^* + B^*$.

(2) $(AB)^* = B^* A^*$.

(3) If α is a complex scalar then $(\alpha A)^* = \overline{\alpha} A^*$.

For example,

$$\Big((2-3i)A \Big)^* = (2+3i)A^*$$

$$(5A)^* = 5A^*$$

$$(2iB)^* = -2iB^*$$

(4) $A^{**} = A$.

(5) If A is invertible then so is A^* and $(A^*)^{-1} = (A^{-1})^*$.

proof of (1)

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

Then

$$A^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \text{etc} \\ \overline{a_{12}} & & \\ & & \end{bmatrix} \quad \text{by definition of conj transpose}$$

$$B^* = \begin{bmatrix} \overline{b_{11}} & \overline{b_{21}} & \text{etc} \\ \overline{b_{12}} & & \\ & & \end{bmatrix}$$

$$A+B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \text{etc} \\ a_{21}+b_{21} & & \\ & & \end{bmatrix}$$

$$(A+B)^* = \begin{bmatrix} \overline{a_{11}+b_{11}} & \overline{a_{21}+b_{21}} & \text{etc} \\ \overline{a_{12}+b_{12}} & & \\ & & \end{bmatrix} \quad \text{by definition of conj transpose}$$

$$= \begin{bmatrix} \overline{a_{11}} + \overline{b_{11}} & \overline{a_{21}} + \overline{b_{21}} & \text{etc} \\ \overline{a_{12}} + \overline{b_{12}} & & \\ & & \end{bmatrix} \quad \text{by algebra of complex numbers (key step)}$$

$$= A^* + B^* \quad \text{QED}$$

proof of (3)

$$\text{Let } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Then

$$(aA)^* = \begin{bmatrix} \overline{\alpha a_{11}} & \cdots & \overline{\alpha a_{m1}} \\ \vdots & & \\ \overline{\alpha a_{1n}} & \cdots & \overline{\alpha a_{mn}} \end{bmatrix} \quad \text{by definition of } \alpha A \text{ and the conjugate trans}$$

$$= \begin{bmatrix} \overline{\alpha}\ \overline{a_{11}} & \cdots & \overline{\alpha}\ \overline{a_{m1}} \\ \vdots & & \\ \overline{\alpha}\ \overline{a_{1n}} & \cdots & \overline{\alpha}\ \overline{a_{mn}} \end{bmatrix} \quad \text{by algebra of complex numbers (key step)}$$

$$= \overline{\alpha}\ A^*$$

determinant of the conjugate transpose

$$|A^*| = \overline{|A|}$$

For example, if det A = 2+ 6i then det A^* = 2 − 6i
(See problem 4 for the proof in the 3 × 3 case.)

Hermitian matrices

> A square matrix H is called Hermitian if $H^* = H$.

This means that H is Hermitian if its diagonal entries are real, and matching entries h_{ij} and h_{ji} are conjugates.
A symmetric matrix (meaning *real* entries with $a_{ij} = a_{ji}$) is a special case of Hermitian.
Here are a few examples and non-examples.

$$\begin{bmatrix} 2 & i & 7 \\ -i & 4 & 3+i \\ 7 & 3-i & 6 \end{bmatrix}$$ Hermitian

$$\begin{bmatrix} 1 & 5 & 6 \\ 5 & 0 & -2 \\ 6 & -2 & 3 \end{bmatrix}$$ symmetric (and Hermitian)

$$\begin{bmatrix} 2 & 5+i \\ 5+i & 3 \end{bmatrix}$$ *not* Herm and *not* symmetric since only *real* matrices are ever called symmetric

$$\begin{bmatrix} 2 & i \\ -i & i \end{bmatrix}$$ *not* Hermitian (diagonal entries aren't real)

> In general, to show that a matrix is Hermitian:
>
> *method 1* Look at the matrix to see if its matching entries are conjugates.
>
> *method 2* To show that say $AB + CD^{-1}$ is Hermitian, given some hypotheses, find $(AB + CD^{-1})^*$ using matrix algebra and show that it simplifies to $AB + CD^{-1}$.
>
> Your argument should look something like this:
>
> $$\begin{aligned}(AB + CD^{-1})^* &= \ldots \\ &= \ldots \\ &= \ldots \\ &= AB + CD^{-1}\end{aligned}$$
>
> Therefore $AB + CD^{-1}$ is Hermitian.

skew Hermitian matrices

> A square matrix K is called skew Hermitian if $K^* = -K$.

This means that K is skew Hermitian if its diagonal entries are pure imaginary, and the matching entries k_{ij} and k_{ji} are of the form a+bi and −a+bi; i.e., $\overline{k_{ij}} = -k_{ji}$

A skew symmetric matrix (meaning *real* entries with $a_{ij} = -a_{ji}$) is a special case of skew Herm.

For example,

$$\begin{bmatrix} 3i & 2-7i & -8+4i \\ -2-7i & -6i & 3 \\ 8+4i & -3 & 0 \end{bmatrix} \text{ is skew Herm}$$

$$\begin{bmatrix} 0 & 3 & 4 \\ -3 & 0 & -5 \\ -4 & 5 & 0 \end{bmatrix} \text{ is skew symmetric (and skew Herm)}$$

In general, to show that a matrix is skew-Hermitian:

method 1 Look at the matrix to see if its matching entries are a+bi and −a+bi.

method 2 To show that say $AB + CD^{-1}$ is skew-Hermitian, given some hypotheses, find $(AB + CD^{-1})^*$ using matrix algebra and show that it simplifies to $-(AB + CD^{-1})$.

Your argument should look something like this:

$$\begin{aligned} (AB + CD^{-1})^* &= \ldots \\ &= \ldots \\ &= \ldots \\ &= -(AB + CD^{-1}) \end{aligned}$$

Therefore $AB + CD^{-1}$ is skew-Hermitian.

example 1

Show that if H is Hermitian then iH is skew Herm.

method 1 (good for 3 × 3's)

Let

$$H = \begin{bmatrix} a & b+ci & d+ei \\ b-ci & f & g+ki \\ d-ei & g-ki & m \end{bmatrix}$$

where all the individual letters except i represent real numbers. Then

$$iH = \begin{bmatrix} ai & -c+bi & -e+di \\ c+bi & fi & -k+gi \\ e+di & k+gi & mi \end{bmatrix}$$

By inspection, iH is skew Herm.

method 2 (good for n × n's in general)

I'll show that $(iH)^* = -iH$.

$$\begin{aligned} (iH)^* &= \bar{i}\, H^* && (* \text{ rule}) \\ &= -i\, H^* && (\text{take conjugate of i}) \\ &= -i\, H && (\text{since H is Herm}) \end{aligned}$$

warning

A Herm matrix must have *real* diagonal entries.

A skew Herm matrix has *pure imaginary* diagonal entries.

Note that 0 is both real and pure imag.

PROBLEMS FOR SECTION 7.2

1. Find A^* if $A = \begin{bmatrix} 2 & -i \\ 3-2i & 4+6i \end{bmatrix}$.

2. Find $|A^*|$ if (a) $|A| = 6 + 2i$ (b) $|A| = 7$ (c) $|A| = -6i$

3. Express the conjugate transpose of A^*BC in terms of A,B,C and their conjugate transposes.

4. Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$ (where a,b,c,...,h,k are complex numbers).

Find $|A^*|$ directly to see that it really is the conjugate of $|A|$.

5. True or False. If H is Herm so is (a) $-H$ (b) H^*

6. Show that if A is square then $A + A^*$ is Herm.

7. If H is Herm and A is square show that A^*HA is also Herm.

8. Show that if a Herm matrix H is invertible then H^{-1} is also Herm.

9. If A is square is there anything special about $A - A^*$. Defend your answer.

10. If K is skew Herm what about K^2 and K^3.

SECTION 7.3 THE VECTOR SPACE C^n

definition of C^n

C^n is the set of n-tuples $(u_1,\dots, u_n)$ of complex numbers.

Addition, subtraction and scalar multiplication (where the scalars are complex numbers) is done componentwise as in R^n.

For example , if

$$u = (1,2,6,3) \text{ and } v = (2,\ 6-3i,\ 7,4)$$

then u and v are in C^4 (u is also in R^4) and

$$u + v = (3,\ 8-3i,\ 13,\ 7)$$

bases for C^n

All of the ideas about bases for R^n carry over to C^n.

The space C^2 is 2-dim with standard basis $\vec{i} = (1,0)$, $\vec{j} = (0,1)$. For example, if

$$u = (2+3i,\ 4-6i)$$

then

$$u = (2+3i)\vec{i} + (4-6i)\vec{j}$$

More generally, C^n is n-dim with the same standard basis as R^n.

the dot product (inner product)

If

$$u = (u_1,\dots, u_n) \quad\text{and}\quad v = (v_1,\dots, v_n)$$

then

$$u\cdot v = \overline{u_1}\ v_1 + \dots + \overline{u_n}\ v_n \qquad \text{(the overbar means conjugate)}$$

For example, if

$$u = (4,\ 2-i,\ i,\ 5) \quad\text{and}\quad v = (2i,\ 3+4i,\ 6,\ 3)$$

then

$$u\cdot v = (4)(2i) + (2+i)(3+4i) + (-i)(6) + (5)(3) = 17 + 13i$$

The dot product in R^n can be considered to be a special case of the dot in C^n. The vectors

$$u = (2,3) \text{ and } v = (4,5)$$

are in R^n as well as in C^n and using the C^n dot product you get

$$u\cdot v = (\overline{2})(4) + (\overline{3})(5) = (2)(4) + (3)(5) = 23,$$

the same as the R^n dot product.

warning

1. Don't forget to conjugate the components of the *first* vector when you compute the dot product in C^n.
2. Some students like to write $u\cdot v = \overline{u}\cdot v$ (with a bar over the u on the righthand side) as a reminder to conjugate the components of the first vector before multiplying and adding. Even if this seems like convenient notation to you, *don't* use it. The dot symbol *includes* the conjugation idea and so the notation $\overline{u}\cdot v$ would suggest that you conjugate *twice*, once because of the bar over the u and once because of the dot symbol (which cancels out to no conjugation at all).

connection between dotting vectors and multiplying matrices

If

$$u = \begin{bmatrix} i \\ 2 \\ 3 \end{bmatrix} \quad \text{and } v = \begin{bmatrix} 4 \\ 5 \\ i \end{bmatrix}$$

then

$$u \cdot v = (-i)(4) + (2)(5) + (3)(i) = 10 - i$$

This procedure is like finding

$$\begin{bmatrix} -i & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ i \end{bmatrix}$$

the product of a row matrix and a col matrix. So for column vectors u and v,

$$\boxed{\text{dot product } u \cdot v \;=\; \text{matrix product } u^* v}$$

properties of the dot product

(1) $v \cdot u = \overline{u \cdot v}$
(2) $(au) \cdot v = \bar{a}(u \cdot v)$
(3) $u \cdot (av) = a(u \cdot v)$
(4) $u \cdot (v + w) = u \cdot v + u \cdot w$

For example, if $u \cdot v = 3i$ then

$v \cdot u = -3i$
$(2u) \cdot v = \bar{2}(u \cdot v) = 2(3i) = 6i$
$(2iu) \cdot v = -2i(u \cdot v) = 6$
$u \cdot (2iv) = 2i(u \cdot v) = -6$

(5) Dot products can be non-real but the dot product of a vector with *itself* must be a non-negative real number and can be 0 only if $u = \vec{0}$. In other words,

$u \cdot u > 0$ if $u \ne \vec{0}$

$u \cdot u = 0$ if $u = \vec{0}$

proof of (1)

Let $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$. Then

$$\begin{aligned} \overline{u\ v} &= \overline{\overline{u_1}\ v_1 + \ldots + \overline{u_n}\ v_n} \\ &= \overline{\overline{u_1}}\ \overline{v_1} + \ldots + \overline{\overline{u_n}}\ \overline{v_n} \quad \text{(conjugate rules)} \\ &= u_1\ \overline{v_1} + \ldots + u_n\ \overline{v_n} \\ &= \overline{v_1}\ u_1 + \ldots + \overline{v_n}\ u_n \\ &= v \cdot u \quad \text{QED} \end{aligned}$$

proof of (2)

Let

$$u = (u_1, \ldots, u_n), \quad v = (v_1, \ldots, v_n)$$

Then

$$
\begin{aligned}
(au)\cdot v &= (au_1,\ldots, au_n)\cdot(v_1,\ldots, v_n) \\
&= \overline{au_1}\, v_1 + \ldots + \overline{au_n}\, v_n && \text{definition of dot} \\
&= \overline{a}\,\overline{u_1}\, v_1 + \ldots + \overline{a}\,\overline{u_n}\, v_n && \text{conjugate rule} \\
&= \overline{a}(\overline{u_1}\, v_1 + \ldots + \overline{u_n}\, v_n) \\
&= \overline{a}(u\cdot v) && \text{definition of dot}
\end{aligned}
$$

orthogonal vectors

> If $\vec{u}\cdot\vec{v} = 0$ then $\vec{u}$ and $\vec{v}$ are called orthogonal.

For example, if $u = (i,1)$ and $v = (-i,1)$ then $u\cdot v = (-i)(-i) + (1)(1) = i^2 + 1 = 0$ so u and v are orthog.

norms

Here is the definition of the norm in several equivalent versions.

> If $u = (u_1,\ldots, u_n)$ then
>
> $$\|u\| = \sqrt{|u_1|^2 + \ldots + |u_n|^2}$$
>
> where $|u_i|$ means the mag of the complex number u_i

> If $u = (u_1,\ldots, u_n)$ then
>
> $$\|u\| = \sqrt{\overline{u_1}\, u_1 + \ldots + \overline{u_n}\, u_n}$$
>
> where $\overline{u_i}$ means the conjugate of the complex number u_i.

> If $u = (a+bi,\ c+di)$ then $\|u\| = \sqrt{a^2 + b^2 + c^2 + d^2}$

The norm in R^n can be considered as a special case of the norm in C^n.

For example, if $v = (2i,\ 3+4i,\ 5)$ then $\|v\| = \sqrt{4 + 9 + 16 + 25} = \sqrt{54}$

warning $\|v\|$ is *not* $\sqrt{(2i)^2 + (3+4i)^2 + 5^2}$

warning

Norms are always real and furthermore are ≥ 0 (as opposed to dot products which can be non-real).

If you end up with a norm of -2 or $5i$ or $\sqrt{3}i$ you made a *mistake*.

properties of norms

(1) $u\cdot u = \|u\|^2$

(2) $\|u\| > 0$ if $u \ne \vec{0}$

$\|u\| = 0$ if $u = \vec{0}$

(3) $\|au\| = |a|\,\|u\|$ where $|a|$ means the mag of the complex number a

For example, if $\|\vec{u}\| = 3$ then

$$\|(2+4i)\vec{u}\| = |2+4i|\,\|\vec{u}\| = \sqrt{20} \times 3 = 3\sqrt{20}$$

$$\|4i\vec{u}\| = |4i|\,\|\vec{u}\| = 4 \times 3 = 12$$

(4) (triangle inequality) $\|u + v\| \le \|u\| + \|v\|$

proof of (1)

Let $\vec{u} = (a_1 + b_1 i, \ldots, a_n + b_n i)$. Then

$$\vec{u}\cdot\vec{u} = (a_1 - b_1 i)(a_1 + b_1 i) + \ldots + (a_n - b_n i)(a_n + b_n i) \quad \text{by dot definition}$$

$$= a_1^2 + b_1^2 + \ldots + a_n^2 + b_n^2$$

$$= \|\vec{u}\|^2$$

proof of (3)

Let $u = (u_1, \ldots, u_n)$ and let a be a scalar. Then $au = (au_1, \ldots, au_n)$ and

$$\|au\| = \sqrt{\overline{au_1}\; au_1 + \ldots + \overline{au_n}\; au_n} \quad \text{by the definition of the norm}$$

$$= \sqrt{\bar{a}\;\overline{u_1}\; au_1 + \ldots + \bar{a}\;\overline{u_n}\; au_n} \quad \text{rule of complex algebra}$$

$$= \sqrt{\bar{a}a(\overline{u_1}\, u_1 + \ldots + \overline{u_n}\, u_n)} \quad \text{ordinary algebra}$$

$$= \underbrace{\sqrt{\bar{a}a}}_{|a|}\;\underbrace{\sqrt{\overline{u_1}\, u_1 + \ldots + \overline{u_n}\, u_n}}_{\|u\|} \quad \text{ordinary algebra}$$

QED

normalized vectors

If $\|\vec{v}\| = 1$ then v is called a unit vector or a normalized vector.
Note that if v is a unit vector then $v\cdot v = 1$ since $v\cdot v = \|v\|^2$.

If u is any vector then the new vector $\dfrac{u}{\|u\|}$ is a unit vector and is called the normalized u. I'll call it u_{unit}.

orthonormal vectors

If $\overrightarrow{u_1}, \ldots, \overrightarrow{u_k}$ are orthogonal unit vectors then they are called orthonormal.

dots and norms in a new coord system

It works as it did in R^n (see §2.3).

If $\overrightarrow{u_1}, \ldots, \overrightarrow{u_n}$ is an *orthonormal* basis for C^n and

$$\vec{x} = X_1\overrightarrow{u_1} + \ldots + X_n\overrightarrow{u_n} = ((X_1, \ldots, X_n))$$

$$\vec{y} = Y_1\overrightarrow{u_1} + \ldots + Y_n\overrightarrow{u_n} = ((Y_1, \ldots, Y_n))$$

then

$$\|x\| = \sqrt{\overline{x_1}\ x_1 + \ldots + \overline{x_n}\ x_n}$$

$$x \cdot y = \overline{x_1}\ y_1 + \ldots + \overline{x_n}\ y_n$$

In other words, the coordinates w.r.t. the basis $\vec{u_1}, \ldots, \vec{u_n}$ (in the double parentheses) can be maneuvered like ordinary $\vec{i}, \vec{j}$ coordinates to find dot products and norms.

coords of a vector w.r.t. an *orthogonal* basis

For the most part, it works as it did in R^n.

If $\vec{u_1}, \ldots, \vec{u_n}$ is an orthogonal basis for C^n and $\vec{x}$ is in C^n then

$$\vec{x} = \frac{u_1 \cdot x}{u_1 \cdot u_1}\vec{u_1} + \ldots + \frac{u_n \cdot x}{u_n \cdot u_n}\vec{u_n}$$

In the even more special case that the basis is ortho*normal*, the formula becomes

$$\vec{x} = (\vec{u_1} \cdot \vec{x})\vec{u_1} + \ldots + (\vec{u_n} \cdot \vec{x})\vec{u_n}$$

The one difference is that in R^n it doesn't matter whether you write the formulas using $u_1 \cdot x$ or $x \cdot u_1$ in the numerator since they are equal. But in C^n you have to use $u_1 \cdot x$ and not $x \cdot u_1$.

PROBLEMS FOR SECTION 7.3

1. If $u = (i,1)$ and $v = (-1, i)$, are u and v orthogonal.

2. Let $u = (2-i, 3, -4i)$, $v = (3-i, -i, 1-2i)$. Find
   iu, $u \cdot v$, $v \cdot u$, $\|u\|$, $\|v\|$, $|u \cdot v|$, v_{unit}

3. Find $\|\vec{u}\|$ if (a) $u = (i,i)$ (b) $u = (2i, 3, -i)$ (c) $u = (3-i, 2+4i)$

4. If If $\vec{u} \cdot \vec{v} = 6-2i$, $\|\vec{u}\| = 3$, $\|\vec{v}\| = 7$ find
   (a) $(\vec{u}+\vec{v}) \cdot (\vec{u}-\vec{v})$
   (b) $\|\vec{u}+i\vec{v}\|$
   (c) $\big((2-3i)\vec{u} + \vec{v}\big) \cdot i\vec{u}$
   (d) $\|6i\vec{u}\|$
   (e) $\|(2-3i)\vec{u}\|$

5. True or False .
   (a) Multiplying a vector by 3 multiplies the norm by 3.
   (b) Multiplying a vector by 3i multiplies the norm by 3i.

6. Look at this statement: If $\|u\| = \|v\|$ then $u + v$ and $u - v$ are orthogonal.
   (a) Use plane geometry to show that it's true in R^2.
   (b) Show that it's true in R^n.
   (c) Show that it isn't true in C^n. First show why the proof from (b) doesn't carry over to C^n and then clinch it with a specific counterexample say in C^2.

7. Show that $\|\vec{x} + i\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 + i(\vec{x} \cdot \vec{y} - \vec{y} \cdot \vec{x})$.

8. Show that $\frac{1}{2}i(\vec{y}\cdot\vec{x}) - \frac{1}{2}i(\vec{x}\cdot\vec{y})$ is the imag part of $\vec{x}\cdot\vec{y}$.
Start by letting $x\cdot y = a + bi$.

9. Let $u = (i,0)$, $v = (0,i)$ be a new basis for C^2.
Let $x = (2+3i,\ 6-7i)$.
Find the new coords of x
(a) by inspection (or by solving some equations)
(b) using the basis changing matrix
(c) using the formula for converting to an orthogonal basis

10. If $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ what does $u_1\overline{v_1} + u_2\overline{v_2} + u_3\overline{v_3}$ compute.

SECTION 7.4 UNITARY MATRICES

unitary matrix rule

Let M be a square matrix. The following are equivalent (either all happen or none happen).

(1) $M^{-1} = M^*$, i.e., $MM^* = I$; i.e., $M^*M = I$.

(2) Columns of M are orthonormal vectors in C^n.

(3) Rows of M are orthonormal vectors in C^n.

A matrix with these properties is called *unitary*.

An orthogonal matrix (meaning *real* entries with $M^{-1} = M^T$) is a special case of unitary.

Note that, for a square matrix, this rule says, among other things:
If the rows are orthonormal then the cols are also orthonormal.
If the cols are orthonormal then the rows are also orthonormal.
If the rows are orthonormal then the inverse of the matrix is easy to find; it's the conjugate transpose.
If the cols are orthonormal then the inverse of the matrix is easy to find; it's the conjugate transpose.

warning

Items (2) and (3) in the unitary matrix rule involve ortho*normal* rows and cols, not just orthogonal rows and cols.

example 1

$$\begin{bmatrix} \frac{1}{\sqrt{10}}\, i & -\frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}}\, i \end{bmatrix}$$ is unitary (easy to see that the cols are orthonormal).

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$ is orthogonal (and unitary).

determinant of a unitary matrix

If U is unitary then $||U|| = 1$, i.e., $|\det U| = 1$.

In other words, the complex number det U has magnitude 1.

proof

$UU^* = I$

$|UU^*| = |I| = 1$ (the vertical bars mean determinant)

$|U|\,|U^*| = 1$

$|U|\ \overline{|U|} = 1$ (since $\det A^* = \overline{\det A}$)

Remember that if z (in this case, det U) is a complex number then $z\overline{z} = |z|^2$.
So $|\det U|^2 = 1$ and since magnitudes are positive, $|\det U| = 1$.

how to show that a matrix is unitary

method 1 Look at its rows (or cols) to see if they are orthonormal.
method 1 Multiply the matrix by its conjugate transpose to see if it is I.

how to show that a matrix is not unitary

method 1 Look at its rows (or cols) to see that they are not orthonormal.
method 2 Show that the matrix times its conjugate transpose is not I.
method 3 Find the det. Then take the magnitude of the determinant. If it isn't 1 then the matrix is not unitary. (But if that mag is 1 then you have no conclusion.)

example 1

If $|M| = 3 + 3i$ then M is not unitary because $||M|| = \sqrt{18}$, not 1.
If $|M| = \frac{1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}$ then $||M|| = 1$ which is inconclusive. M might be unitary.

warning

Note that in the notation $||M||$, the inner vertical bars mean determinant which comes out to be a complex number (possibly real) and the outer vertical bars mean the magnitude of that number.
Instead of $||M||$ you can write $|\det M|$ or you can write "mag of det M".

The notation $||M||$ doesn't mean norm because we never defined anything called the norm of a matrix.

warning

If U is unitary, the rule is *not* $|U| = 1$ and it is *not* $|U| = \pm 1$.

The rule is $||U|| = 1$; i.e., mag of det U is 1.

example 2

Suppose M is unitary. To show that iM is also unitary, I'll show that $(iM)(iM)^* = I$:

$$\begin{aligned}
(iM)(iM)^* &= iM\,\bar{i}\,M^* && \text{by } * \text{ rule} \\
&= -i^2\, MM^* && \text{matrix algebra} \\
&= MM^* \\
&= I && \text{since M is unitary}
\end{aligned}$$

PROBLEMS FOR SECTION 7.4

1. Show that the product of unitary matrices is unitary.

2. Show that the transpose of a unitary matrix is unitary.

3. Can you tell whether or not M is unitary if $|M|$ is
(a) 2 (b) -1 (c) $1 + 2i$ (d) $\frac{1}{2} + \frac{1}{2}i\sqrt{3}$
(e) $-i$ (f) $\frac{3}{5} - \frac{4}{5}i$

4. If U is unitary, what about $-3U$.

5. If U is unitary show that U^{-1} is unitary.

6. Show that the sum of unitary matrices is not necessarily unitary.

REVIEW PROBLEMS FOR CHAPTER 7

1. Let $u = (2i, -6)$, $v = (3+4i, 5+6i)$. Find $u\cdot v$, $v\cdot u$, $\|u\|$, $\|v\|$.

2. If $\|w\| = 3$, simplify $\Big(v - (w\cdot v)w\Big)\cdot w$.

3. Show that if H is Hermitian then $|H|$ is real.

4. Show that if H is Hermitian and U is unitary then $U^{-1}HU$ is Hermitian.

5. Prove or disprove: The product of Hermitian matrices is Hermitian.

6. Let

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & -i \end{bmatrix}$$

(a) Show that A is unitary.
(b) Find A^{-1}.

7. Let

$$A = \begin{bmatrix} 0 & x \\ 3i & y \end{bmatrix}$$

Find x and y if possible so that A is

(a) unitary
(b) Hermitian
(c) skew Hermitian

8. Find u_{unit} if $u = (i, i, 2, 3+2i)$

9. Let u,v,w be in C^n. Let $A = \begin{bmatrix} u\cdot u & u\cdot v & u\cdot w \\ v\cdot u & v\cdot v & v\cdot w \\ w\cdot u & w\cdot v & w\cdot w \end{bmatrix}$.

What is special about the matrix A.

CHAPTER 8 EIGENVALUES AND EIGENVECTORS

SECTION 8.1 INTRODUCTION

definition of eigenvalue and eigenvector

Let M be a square matrix. Suppose $\vec{x}$ is a nonzero column vector, λ is a scalar and

$$M\vec{x} = \lambda\vec{x}$$

Then λ is called an eigenvalue of M and $\vec{x}$ is the corresponding eigenvector.
In other words, an eigenvector of M is a nonzero vector that maps to a multiple of itself.

Note that $M\vec{0} = \vec{0}$ so $\vec{0}$ always maps to a multiple of itself, but $\vec{0}$ doesn't count as an eigenvector because eigenvectors must be nonzero.

eigenspaces

Suppose u and v are eigenvectors of an $n \times n$ matrix corresponding to the eigenvalue λ. Let k be a scalar.
Then u + v and ku are also eigenvectors corresponding to λ, provided that they are not $\vec{0}$ (i.e., provided $v \neq -u$ and $k \neq 0$).
The set of all eigenvectors corresponding to eigenvalue λ, together with the zero vector, is a subspace of R^n or C^n called the eigenspace corresponding to λ.

proof

Let u and v be eigenvectors corresponding to eigenvalue λ.
I want to show that if u + v and au are nonzero then they are also eigenvectors corresponding to λ.

$M(u+v)$	$= Mu + Mv$	matrix algebra
	$= \lambda u + \lambda v$	u and v are eigenvector corr to λ
	$= \lambda(u + v)$	vector algebra

So provided that u+v is not $\vec{0}$, it is an eigenvector corr to λ.
Also

$M(ku)$	$= k(Mu)$	matrix algebra
	$= k(\lambda u)$	u is an eigenvector corr to eigenvalue λ
	$= \lambda(ku)$	vector algebra

So provided that ku is not $\vec{0}$, it's an eigenvector corr to λ.

the awkward status of the zero vector

It turns out that *non*zero vectors with the property that $M\vec{x} = \lambda\vec{x}$ are useful.

But the fact that $\vec{0}$ always has this property is never interesting or useful which is why mathematicians do not consider the vector $\vec{0}$ to be an eigenvector of a matrix M.
This unfortunately means that the set of all eigenvectors of M corresponding to an eigenvalue λ is not a subspace because a subspace must contain $\vec{0}$.
So the eigenspace of M corresponding to eigenvalue λ is defined as the set of all eigenvectors corresponding to λ plus the non-eigenvector $\vec{0}$.

example

Suppose M reflects points in the x-axis in R^2. (Fig 1).
If u is on the x-axis then Mu = u so u is an eigenvector corresponding to eigenvalue 1.
If v is on the y-axis then Mv = $-$v so v is an eigenvector corresponding to eigenvalue -1.
If w is on neither axis then Mw is not a multiple of w, so w is not an eigenvector.
The eigenvalues are 1 and -1. The corresponding eigenspaces are the x-axis and y-axis respectively.

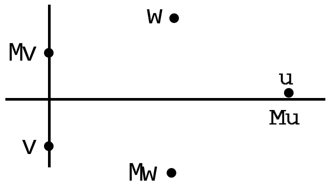


FIG 1

mathematical catechism

question 1 What does it mean to say that A has eigenvalue λ.
answer It means that there is a nonzero vector u such that $Au = \lambda u$.

question 2 What does it mean to say that A has eigenvector u.
answer It means that $u \ne \vec{0}$ and there is a scalar λ such that $Au = \lambda u$.

question 3 What does it mean to say that A has eigenvector u with corresponding eigenvalue λ.
answer It means that $u \ne \vec{0}$ and $Au = \lambda u$.

PROBLEMS FOR SECTION 8.1

1. Let

$$A = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad u = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Are u and v eigenvectors of A? If so, find the corresponding eigenvalues.

2. Let u be a nonzero vector. Spot eigenvalues and eigenvectors if

(1) $Mu = 6u$ (2) $Mu = \vec{0}$ (3) $ABu = 3u$ (4) $ABu = 3Bu$
(5) $ABu = 3Au$ (6) $2Mu = u$ (7) $M(u+v) = 3u + 3v$ (8) $(A + B)u = Au + Bu$
(9) $(A+B)u = 2u$

3. Suppose u is an eigenvector of M corresponding to eigenvalue λ.
Show that u is also an eigenvector of 3M and find the corresponding eigenvalue.

4. Find the eigenvalues and eigenvectors of the $n \times n$ matrix I.

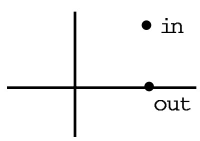
5. Remember that the null space of a matrix operator M is the set of vectors that map to $\vec{0}$.

(a) If the null space contains more than just $\vec{0}$ show that it is an eigenspace of M and find the corresponding eigenvalue.

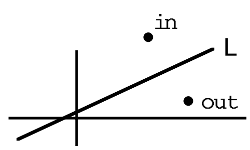
(b) By the way, what kind of matrix has a null space that contains more than $\vec{0}$.

6. Suppose u is an eigenvector of matrix M corresponding to eigenvalue $\lambda = 5$. What can you conclude about u_{unit}.

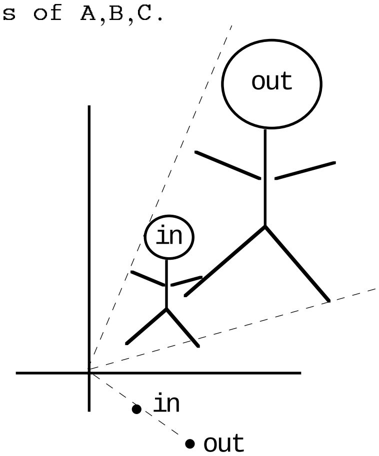
7. Suppose A,B,C are 2×2 matrices.
A projects points onto the x-axis as illustrated in the first diagram.
B reflects points in a line L through the origin as illustrated in the middle diagram.
C expands radially by a factor of 2 as illustrated in the third diagram.
Think geometrically to find eigenvalues and eigenspaces of A,B,C.



Problem 7A



Problem 7B



Problem 7C

8. (a) Can an eigenvalue of a matrix M have more than one corresponding eigenvector.
(b) Can an eigenvector of a matrix M have more than one corresponding eigenvalue.

9. Let A be square with eigenvalue λ and corresponding eigenvector u.
Let $B = P^{-1}AP$.
Show that B also has eigenvalue λ and find the corresponding eigenvector.

10. Can two eigenspaces (corresponding to different eigenvalues) overlap (i.e., have something in common)?

11. Suppose 0 is an eigenvalue of M. What does this have to do with invertibility.
Suggestion: Remember what it means for 0 to be an eigenvalue and then use the invertible matrix rule.

SECTION 8.2 FINDING EIGENVALUES AND EIGENVECTORS

finding eigenvalues

(1)

> To find eigenvalues of an $n \times n$ matrix M, solve this equation for λ:
>
> $$|M - \lambda I| = 0$$
>
> i.e., find λ's that make the determinant of the matrix $M - \lambda I$ zero.

Here's why.

To find eigenvalues of M you want to find λ's so that the equation $M\vec{x} = \lambda\vec{x}$ has a *nonzero* solution for $\vec{x}$. The equation can be written as

$$(M - \lambda I)\vec{x} = \vec{0}$$

It's a square system and by Cramer's rule it has just one solution iff $|M - \lambda I| \neq 0$.

It also is a homog system and always has (at least) the trivial solution $\vec{x} = \vec{0}$. To get a *non*zero solution in addition to the trivial solution you must *avoid* having just one solution, i.e., you must avoid $|M-\lambda I| \neq 0$. So you want $|M-\lambda I| = 0$. QED

characteristic polynomial and characteristic equation

If M is $n \times n$ then $|M-\lambda I|$ is an n-th degree polynomial in the variable λ (you'll see this when you start doing examples), called the *characteristic polynomial* of M, and $|M-\lambda I| = 0$ is called the *characteristic equation*.

the multiplicity of an eigenvalue

Suppose the characteristic equation factors into

$$(\lambda-8)^4(\lambda-6)^3(\lambda-7) = 0$$

Then $\lambda=8$ is called a 4-fold root of the equation, or a *4-fold eigenvalue* or an eigenvalue with *multiplicity* 4. Similarly 6 is a 3-fold eigenvalue and 7 is a 1-fold eigenvalue.

how many eigenvalues does a matrix have

(2)

> An $n \times n$ matrix has at least one and at most n eigenvalues, and their multiplicities add up to n. We say that the matrix has n eigenvalues counting multiplicity meaning that a 3-fold eigenvalue counts 3 times.

This holds because the characteristic equation is an n-th degree polynomial equation in variable λ and that kind of an equation always has n (possibly non-real) roots counting multiplicity.

For example, if M is 6×6 then M can have one eigenvalue with multiplicity 6.
Or M can have one 5-fold eigenvalue and one 1-fold eigenvalue.
Or M can have six 1-fold eigenvalues, etc.

finding eigenvectors after using (1) to find eigenvalues

If λ is an eigenvalue of M then the corresponding eigenvectors are found by solving $M\vec{x} = \lambda\vec{x}$ for $\vec{x}$. In other words:

> The eigenspace corresponding to eigenvalue λ is the set of solutions to the homog system
>
> $$(M - \lambda I)\vec{x} = \vec{0}.$$

Furthermore:

(2)

> If λ is a k-fold eigenvalue of M then there are k or fewer corresponding independent eigenvectors (proof omitted).
>
> In particular, a 1-fold eigenvalue has one corresponding ind eigenvector; a 3-fold eigenvalue may have 1, 2 or 3 corresponding ind eigenvectors, etc.

example 1

Let

$$M = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}$$

To find the eigenvalues and a maximal number of independent eigenvectors for each eigenvalue, begin with

$$M - \lambda I = \begin{bmatrix} 5-\lambda & -2 & 4 \\ -2 & 8-\lambda & 2 \\ 4 & 2 & 5-\lambda \end{bmatrix}$$

The characteristic poly is

$$|M-\lambda I| = (5-\lambda)\Big[(8-\lambda)(5-\lambda) - 4\Big] + 2\Big[-2(5-\lambda) - 8\Big] + 4\Big[-4 - 4(8-\lambda)\Big]$$

(I expanded the det across row 1)

$$= (5-\lambda)\Big[\lambda^2 - 13\lambda + 36\Big] + 4[\lambda-9] + 16[\lambda-9]$$
$$= (5-\lambda)(\lambda-4)(\lambda-9) + 20(\lambda-9)$$
$$= (\lambda-9)\Big[(5-\lambda)(\lambda-4) + 20\Big]$$
$$= (\lambda-9)(-\lambda^2 + 9\lambda)$$
$$= -\lambda(\lambda-9)^2$$

The characteristic equ is $-\lambda(\lambda-9)^2 = 0$ and the roots are $\lambda = 0,9,9$.
0 is 1-fold so there is one corresponding eigenvector.
9 is 2-fold so there are either one or two corresponding independent eigenvectors.

case where $\lambda = 0$

To find the eigenvectors, solve $(M - \lambda I)\vec{x} = \vec{0}$ for $\vec{x}$ which in this case means solving $M\vec{x} = \vec{0}$. Start with

$$\left[\begin{array}{ccc|c} 5 & -2 & 4 & 0 \\ -2 & 8 & 2 & 0 \\ 4 & 2 & 5 & 0 \end{array}\right]$$

and row op to get

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Solution is

$$x_2 = -\frac{1}{2}x_3$$
$$x_1 = -x_3$$

Choose any one solution. If $x_3 = -2$ then $x_2 = 1$, $x_1 = 2$ so an eigenvector is $u = (2,1,-2)$. (All other eigenvectors corresponding to $\lambda=0$ are multiples of u.)

case where $\lambda = 9$

Solve $(M - 9I)\vec{x} = \vec{0}$, i.e., solve

$$\begin{array}{ccc|c} -4 & -2 & 4 & 0 \\ -2 & -1 & 2 & 0 \\ 4 & 2 & -4 & 0 \end{array}$$

Each equation says $-2x_1 - x_2 + 2x_3 = 0$.
Use say x_1 and x_3 as free variables; then $x_2 = 2x_3 - 2x_1$.
Since there are two free variables there are two ind eigenvectors.
Set $x_1 = 1$, $x_3 = 0$ to get eigenvector $v = (1,-2,0)$.
Set $x_1 = 0$, $x_3 = 1$ to get eigenvector $w = (0,2,1)$.

example 2

Let

$$M = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Find eigenvalues, and for each λ, find as many ind eigenvectors as possible.

solution $|M-\lambda I| = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$

The sols to $\lambda^2 + 1 = 0$ are $\lambda = \pm i$; each λ is 1-fold.

To find the eigenvectors corresponding to $\lambda=i$, solve $(M-iI)\vec{x} = \vec{0}$,

$$\begin{array}{cc|c} -i & -1 & 0 \\ 1 & -i & 0 \end{array}$$

Both equations say $x = iy$.
Set $y = 1$ to get eigenvector $u = (i,1)$

To find the eigenvectors corresponding to $\lambda = -i$, solve $(M + iI)\vec{x} = \vec{0}$,

$$\begin{array}{cc|c} i & -1 & 0 \\ 1 & i & 0 \end{array}$$

The first equation says $x = y/i$. The second says $x = -iy$. since $1/i = -i$, these are the same equation. Set $y = 1$ to get eigenvector $v = (-i,1)$.

Note that a real matrix may have non-real eigenvalues in which case the eigenvectors will be non-real also.

matrices with a complete set of eigenvectors

If an $n \times n$ matrix M has n ind eigenvectors then M is said to have a complete set of eigenvectors, namely, enough eigenvectors to make a basis for C^n (or, if real, for R^n).

> This happens iff each k-fold eigenvalue of M has k corresponding ind eigenvectors.
>
> As a special case, if all the eigenvalues of M are 1-fold then M must have a complete set of eigenvectors.

Here's how to get the n independent eigenvectors when they do exist.
Suppose M is 6×6 with

3-fold eigenvalue λ_1
2-fold eigenvalue λ_2
1-fold eigenvalue λ_3.

And suppose that each k-fold λ produces k ind eigenvectors. Then there are three ind eigenvectors u,v,w corresponding to λ_1; there are two ind eigenvectors p,q corresponding to λ_2; and there is one ind eigenvector r corresponding to λ_3. It can be shown (proof omitted here) that the six eigenvectors u,v,w,p,q,r are ind.

The matrix M in example 1 has a 1-fold eigenvalue 0 with eigenvector u = (2,1,−2) and a 2-fold eigenvalue 9 with ind eigenvectors v = (1,−2,0), w = (0,2,1). So M has a complete set of eigenvectors, namely u,v,w, available to be a basis for R^3.

eigenvalues and eigenvectors of a diagonal matrix

Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. By inspection,

$$Ai = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = 2i, \quad Aj = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = 2j, \quad Ak = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} = 3k.$$

So 2 is an eigenvalue with corresponding eigenvectors i,j (a 2-dim eigenspace) and 3 is an eigenvalue with corresponding eigenvector k (a 1-dim eigenspace).

invertible rule

Let M be $n \times n$.

The following are equivalent; i.e., either all are true or all are false.

(1) M is invertible (nonsingular).

(2) $|M| \neq 0$.

(3) Echelon form of M is I.

(4) Rows of M are independent.

(5) Cols of M are independent.

(6) Rank of M is n.

(7) $Mx = \vec{0}$ has only the trivial solution $x = \vec{0}$.

In other words, the null space of M contains only $\vec{0}$.

In other words, if $x \neq \vec{0}$ then $Mx \neq \vec{0}$.

In other words, the operator M maps $\vec{0}$, *and nothing else*, to $\vec{0}$.

(8) The operator M is one-to-one, i.e., M *doesn't* send two inputs to the same output.

(9) The range of M is R^n.

(10) The eigenvalues of M are all nonzero.

Having (10) on the list means every *non*-invertible matrix has 0 as an eigenvalue and *no* invertible matrix has 0 as an eigenvalue.

proof that (10) belongs on the list

I'll show that *not* having (10) is equivalent to *not* having (2).

0 is an eigenvalue of M iff the equation $|M - \lambda I| = 0$ has $\lambda = 0$ as one of its sols
iff $|M - 0I| = 0$
iff $|M| = 0$

leading term of the characteristic poly

If M is 3×3 then the characteristic poly of M begins with $-\lambda^3$ (as in example 1).

If M is 4×4 then the characteristic poly of M begins with λ^4.

In general suppose M is $n \times n$.
If n is even then the characteristic poly of M begins with λ^n.
If n is odd then the characteristic poly of M begins with $-\lambda^n$.

proof.

Let $M = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$

Then

$$|M - \lambda I| = \begin{vmatrix} a_{11}-\lambda & \dots & a_{1n} \\ \vdots & & \\ a_{n1} & \dots & a_{nn}-\lambda \end{vmatrix}$$

Compute the determinant by finding the sum of products consisting of one entry from each row and col where each product in the sum is prefixed a sign according to the how-many-inversions rule.

One of the products in the sum is

$$(a_{11} - \lambda)(a_{22} - \lambda) \;\dots\; (a_{nn} - \lambda),$$

the product consisting of all the diagonal entries of $A-\lambda I$. This product always gets a plus sign since there are no inversions. The λ^n term in the sum comes entirely from multiplying out this product. And if there are an even number of factors, i.e., if n is even, then you get λ^n when you multiply out; if n is odd, you get $-\lambda^n$.

sum and product of the eigenvalues

(1) The product of the eigenvalues of M is $|M|$ provided that a k-fold λ is counted k times in the product. For example, if the eigenvalues are $-2,3,3$ then $|M| = -18$.
(2) The trace of a square matrix M is defined as the sum of its diagonal entries.
The sum of the eigenvalues of M is trace M provided that a k-fold λ is counted k times in the sum.

For example if $M = \begin{bmatrix} 2 & \cdot \\ \cdot & 4 \end{bmatrix}$ then trace M = 6 and the sum of the eigenvalues must be 6.

proof in the 3×3 case

Let

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$$

Suppose M has eigenvalues $\lambda_1, \lambda_2, \lambda_3$.

I'm going to compute the characteristic poly $|M - \lambda I|$ in two different ways.

For the first way, expand the determinant using the original definition of a determinant in Section 1.3.

$$(*) \quad |M - \lambda I| = \begin{vmatrix} a-\lambda & b & c \\ d & e-\lambda & f \\ g & h & k-\lambda \end{vmatrix} = (a-\lambda)(e-\lambda)(k-\lambda) - (a-\lambda)hf + \text{four more terms}$$

For the second way, remember that $\lambda_1, \lambda_2, \lambda_3$ are roots of the characteristic poly so the polynomial has factors $\lambda-\lambda_1$, $\lambda-\lambda_2$, $\lambda-\lambda_3$. And

$$(**) \quad |M - \lambda I| = -(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$$

The minus sign is there because if M is $n \times n$ where n is *odd* then the leading coeff of the characteristic poly is $-\lambda^n$.

The two versions in (*) and (**) must agree.

Equate coeffs of λ^2 from the two versions in (*) and (**): The λ^2 coeff in (*) is a+e+k. The λ^2 coeff in (**) is $\lambda_1+\lambda_2+\lambda_3$. So $\lambda_1+\lambda_2+\lambda_3 = a + e + k$, proving (2).

Equate the constant terms in (*) and (**): The constant term in (**) is $\lambda_1\lambda_2\lambda_3$. The constant term in (*) includes aek − ahf plus four more terms. Instead of doing algebra to get it, notice that for any polynomial you can get the constant term by setting the variable equal to 0. So to get the constant term in (*), set $\lambda=0$ in $|M-\lambda I|$. The constant term is $|M|$. So $|M| = \lambda_1\lambda_2\lambda_3$, proving (1).

mathematical catechism

question 1 What does it mean to say that a square matrix has a complete set of eigenvectors.

first answer It means that if M is $n \times n$ then M has n independent eigenvectors.

second answer It means that each k-fold eigenvalue has k corresponding independent eigenvectors.

question 2 What is the characteristic polynomial of a square matrix M.

answer $|M - \lambda I|$

question 3 What is the characteristic equation of a square matrix M.

answer $|M - \lambda I| = 0$

question 4 What does it mean to say that $\lambda=7$ is a 3-fold eigenvalue of A.

answer It means that $(\lambda-7)^3$ is a factor of the characteristic poly $|A - \lambda I|$.

PROBLEMS FOR SECTION 8.2

1. Find eigenvalues and corresponding independent eigenvectors.
Does the matrix have a complete set of eigenvectors.

(a) $\begin{bmatrix} 0 & 3 \\ 2 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 2 & -6 \\ 2 & -1 & -3 \\ -2 & -1 & 1 \end{bmatrix}$ (the eigenvalues turn out to be $\lambda = -2,-2,6$)

(d) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 4 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & \pi \end{bmatrix}$

2. The matrix that rotates points in R^2 by θ degrees counterclockwise is

$$M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

(a) Find the eigenvalues (the algebra comes out nicer than you would expect).
(b) Pick one of your eigenvalues and find the corresponding eigenvector(s).

3. Look at example 1 where I found eigenvectors u,v,w. Find a dozen other eigenvectors.

4. Suppose M is 6×6 and λ is one of its eigenvalues. If u and v are eigenvectors corr to λ and u and v happen to be ind, what can you conclude about the multiplicity of λ.

5. Suppose M is 5×5. Is it possible for M to not have eigenvalues or eigenvectors?

6. Suppose

A has characteristic poly $(\lambda-3)^5(\lambda-2)^2(\lambda-4)$

B has characteristic poly $-(\lambda+1)(\lambda-2)(\lambda+3)$.

Find the size of each matrix.
Find their eigenvalues.
What can you conclude about the number of ind eigenvectors for each eigenvalue.
Do the matrices have a complete set of eigenvalues.

7. (a) Suppose the book's answer says that the eigenvector of M corresponding to the 1-fold eigenvalue λ is $(i,1)$ but your answer is $(-1,i)$. Are you a dope?
(b) Suppose the book's answer says that the eigenvectors of M corresponding to the 2-fold eigenvalue λ are $u = (1,1,1,1)$, $v = ((2,0,3,1)$. Your eigenvectors are $u = (1,1,1,1)$ (so far so good) and $p = (1,2,1,1)$. Are you right.

8. What can you conclude about eigenstuff if
(a) $|A - 2I| = 0$ (b) $|A - 2I| = 6$

9. Find eigenvalues and eigenvectors for
(a) the 3×3 identity matrix I (b) the 3×3 zero matrix
Find them by inspection and then, for practice, with the method used in example 1 (although that is overkill).

10. (a) Factor $P^{-1}AP - \lambda I$.
(b) Use part (a) to show that similar matrices have the same characteristic polynomials
(c) Show that similar matrices have the same eigenvalues.

11. Suppose A is 5×5, 9 is an eigenvalue, and $A - 9I$ has rank 2.
(a) How many independent eigenvectors are there corresponding to eigenvalue 9.
(b) How many total eigenvectors are there corresponding to eigenvalue 9.

12.(a) When is 0 an eigenvalue of a matrix M.
(b) If 0 *is* an eigenvalue of an $n \times n$ matrix M with rank r, what is the dimension of the corresponding eigenspace?

13. Suppose M is invertible and has eigenvector u with corresponding eigenvalue λ. Find some eigenstuff for M^{-1}.

14. Suppose the matrix AB has eigenvalue 5. What determinant must therefore be 0.

15. Let $A = \begin{bmatrix} 0 & \pi & 0 \\ 3 & \sqrt{2} & 7 \\ 1 & 2 & 1 \end{bmatrix}$.

(a) Find the product of the eigenvalues of A.
(b) Find the sum of the eigenvalues of A.

16. (a) True or False. If u,v,w are ind vectors in R^5 and p,q are also ind vectors in R^5 then u,v,w,p,q are ind.

(b) True or False. If u,v,w are ind eigenvectors of M corresponding to $\lambda=3$, and p,q are ind eigenvectors of M corresponding to $\lambda=\pi$ then u,v,w,p,q are ind.

17. On an exam, a student was given a 2×2 matrix and told to find eigenvalues and eigenvectors.

She found the eigenvalues correctly.

Then she tried to solve the system of equations $(M - \lambda I)\vec{x} = \vec{0}$ to find the eigenvectors corresponding to the first eigenvalue λ. After doing some algebra she ended up writing "the only solution I can find is $x_1=0$, $x_2=0$ and since (0,0) does not count as an eigenvector, this λ doesn't have any corresponding eigenvectors".

Could she be right. Could some mean teacher have made up an exam question like that.

EIGENVALUES AND EIGENVECTORS OF HERMITIAN MATRICES

lemma (sliding property of Herms)

If H is an $n \times n$ Hermitian matrix and u and v are in C^n and written as column vectors then

$$Hu \cdot v = u \cdot Hv$$

proof

$Hu \cdot v = (Hu)^* v$ the dot product $p \cdot q$ is the same as the matrix product $p^* q$ (§7.3)
$= u^* H^* v$ by * rule
$= u^* Hv$ since H is Herm
$= u \cdot Hv$ Section 7.3 again

properties of eigenvalues and eigenvectors of Hermitian (and symmetric) matrices

Let H be an $n \times n$ Hermitian matrix.
(1) The eigenvalues are real.
(2) Eigenvectors corresponding to different eigenvalues are orthogonal.
(3) Each k-fold eigenvalue has k corresponding ind eigenvectors.
(4) H has a complete set of orthonormal eigenvectors (i.e., has n of them).

And here's how to find the complete set.

Suppose H is 4×4 with 1-fold eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. Pick corresponding eigenvectors $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4$. They will automatically be orthog by property (2). Normalize them and they will still be eigenvectors corresponding to $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ (multiples of eigenvectors are still eigenvectors) and still be orthog. So $\vec{u}_{1\text{unit}}, \vec{u}_{2\text{unit}}, \vec{u}_{3\text{unit}}, \vec{u}_{4\text{unit}}$ are a complete set of orthonormal eigenvectors for H.

Or suppose H is 4×4 with 3-fold eigenvalue λ_1 and 1-fold eigenvalue λ_2. Pick 3 independent eigenvectors $\vec{x}_1, \vec{x}_2, \vec{x}_3$ corresponding to λ_1. Use the Gram Schmidt process to exchange them for 3 *orthog* vectors $\vec{u}_1, \vec{u}_2, \vec{u}_3$ in the eigenspace. Pick any eigenvector $\vec{w}$ corresponding to λ_2. By (2), $\vec{w}$ is orthog to $\vec{u}_1, \vec{u}_2, \vec{u}_3$. Normalize $\vec{w}$, $\vec{u}_1, \vec{u}_2, \vec{u}_3$ and they will be a complete set of orthonormal eigenvectors of H.

proof of (1)

Let λ be an eigenvalue of H. I want to show that λ is real.
There is a nonzero vector x such that $Hx = \lambda x$. Then

$x \cdot Hx = Hx \cdot x$ (by the sliding property of Herms)
$x \cdot \lambda x = \lambda x \cdot x$ ($Hx = \lambda x$)
$\lambda(x \cdot x) = \overline{\lambda}(x \cdot x)$ (dot rules)
$\lambda = \overline{\lambda}$ (it's OK to cancel the $x \cdot x$'s since $x \neq \vec{0}$ so $x \cdot x \neq 0$)

So λ is real since it equals its conjugate.

proof of (2)

Let $\vec{u}_1$ and $\vec{u}_2$ be eigenvectors corresponding to different eigenvalues λ_1, λ_2. Write them as column vectors. Then

$u_1 \cdot Hu_2 = Hu_1 \cdot u_2$ (sliding property)

$u_1 \cdot (\lambda_2 u_2) = \lambda_1 (u_1 \cdot u_2)$ ($Hu_1 = \lambda_1 u_1$, $Hu_2 = \lambda_2 u_2$)

$\lambda_2 (u_1 \cdot u_2) = \lambda_1 (u_1 \cdot u_2)$ (dot rules plus the fact that, by (1), λ_1 is real)

$(\lambda_2 - \lambda_1)(u_1 \cdot u_2) = 0$

So $\lambda_2 = \lambda_1$ or $u_1 \cdot u_2 = 0$

But λ_2 and λ_1 are different eigenvalues so it must be that $u_1 \cdot u_2 = 0$.

So u_1 and u_2 are orthogonal, QED

proof of (3) too hard

warning

Herms are not the only matrices with properties (1), (2), (3) or (4). But it can be shown that they are the only ones with all four properties simultaneously.

PROBLEMS FOR SECTION 8.3

1. Suppose A is Herm and has eigenvalue 2 with corresponding eigenvector u = (1,3) and also has eigenvalue 3 with corresponding eigenvector v = (2,y). Find y.

2. Suppose that among the eigenvectors of a matrix M are u = (2,3) and v = (4,5). Can you tell if M is Hermitian.

3. Can you tell if M is Hermitian if its eigenvalues are
(a) 2,3,4 (b) $2 \pm 4i$, 6

SECTION 8.4 DIAGONALIZING A SQUARE MATRIX

the diagonalizing process

Diagonalizing a square matrix A means finding a matrix P (if possible) so that $P^{-1}AP$ is diagonal.

Here's how to do it.

Suppose a 3×3 matrix A has the three independent eigenvectors

$$u = (u_1, u_2, u_3), \quad v = (v_1, v_2, v_3), \quad w = (w_1, w_2, w_3)$$

with corresponding eigenvalues $\lambda_1, \lambda_2, \lambda_3$ (not necessarily distinct, e.g., could have $\lambda_2 = \lambda_3$). Let

$$P = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} \quad \text{(the basis changing matrix)}$$

Then

$$(1) \qquad P^{-1}AP = \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}}_{\text{call this } \Lambda}$$

We say that P *diagonalizes* A.

This means that if a transformation is represented by matrix A w.r.t. the usual basis i,j,k then with respect to a a new basis consisting of eigenvectors of A, the transformation is represented by the diagonal matrix Λ.

The process of diagonalizing A is also referred to as finding a diagonal matrix *similar* to A (similar matrices were defined in Section 6.2).

Furthermore it can be shown that (1) is the *only* way to diagonalize A, meaning that if say $Z^{-1}AZ = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix}$ then 2,5,7 are eigenvalues of A and the columns of Z are the corresponding eigenvectors.

when is a matrix diagonalizable

The matrix A can be diagonalized iff it has a complete set of eigenvectors.
Remember that this happens iff each k-fold λ has k independent eigenvectors.
And in particular, it happens if all the eigenvalues are 1-fold.

A matrix that doesn't have "enough" eigenvectors can't be diagonalized.

proof of (1)

Look at the underlying transformation T, the transformation represented by matrix A w.r.t. basis i,j,k and represented by $P^{-1}AP$ w.r.t. the basis u,v,w. I want to show that $P^{-1}AP$ is Λ.

You know that

$$T(u) = \lambda_1 u$$

With respect to basis u,v,w, the coords of u are 1,0,0 and the coords of $\lambda_1 u$ are λ_1,0,0. So

$$P^{-1}AP \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ 0 \\ 0 \end{bmatrix} \tag{2}$$

Similarly,

$$P^{-1}AP \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda_2 \\ 0 \end{bmatrix} \quad \text{and} \quad P^{-1}AP \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \lambda_3 \end{bmatrix} \tag{3}$$

The three equations in (2) and (3) force $P^{-1}AP$ to be Λ.

diagonalizing Hermitian (and symmetric) matrices

Suppose H is Hermitian.
We know from the last section that H has a complete set of eigenvectors so H is diagonalizable. Furthermore we know that H has a complete set of *orthonormal* eigenvectors so it is always possible to choose a *unitary* P for the diagonalizing process. And we know that H has real eigenvalues so Λ is real. All in all:

> If H is Herm then H is diagonalizable; i.e., you can find a P so that $P^{-1}\Omega HP$ = Ò. Furthermore Ò is real and P can be chosen to be unitary (if you play your cards right).

Here's a reminder of how to get a unitary P. For each 1-fold eigenvalue of H, pick an eigenvector. For each k-fold eigenvalue of H, $k > 1$, pick k ind eigenvectors and Gram Schmidt them. This produces a complete set of orthogonal eigenvectors. Normalize them (they're still eigenvectors) and use them as cols of P.

warning

1. Herms are not the only diagonalizable matrices. They are not even the only matrices which can be diagonalized with a unitary P. You may be lucky enough to get a complete set of eigenvectors (maybe even orthogonal ones) for a non-Herm.
  But Herms are the only matrices that can be diagonalized so that P is unitary *and* Λ is real (proof in problem 7).

2. If you want to make P a unitary matrix (in the real case, an orthogonal matrix) don't forget to normalize the orthogonal eigenvectors.
  But there's no point normalizing non-orthog eigenvectors so don't go on a normalizing binge.

example 1

Let

$$A = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}$$

Diagonalize A if possible.

solution A is symmetric so it is diagonalizable. To get a diagonalization, first find the eigenvalues and a complete set of eigenvectors.
  From example 1 in the preceding section, A has these eigenvalues and eigenvectors:
  $\lambda_1 = 0$ (1-fold) with corresponding eigenvector $u = (2,1,-2)$
  $\lambda_2 = 9$ (2-fold) with corresponding eigenvectors $v = (0,2,1)$, $w = (1,-2,0)$.

There are a variety of diagonalizations possible, depending on which eigenvectors are used and on the order in which they are lined up as cols of P.
If the cols of P are u,v,w in that order then the diagonal entries in Λ are 0,9,9 in that order. So one answer to the question is

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} \text{ where } P = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & -2 \\ -2 & 1 & 0 \end{bmatrix}$$

As a check, you can actually compute $P^{-1}AP$ to see that it is diagonal but it's not necessary to do that. If your algebra is correct then $P^{-1}AP$ is guaranteed to be the diagonal matrix Λ.

Similarly, if the cols of P are v,u,w in that order then the diagonal entries in Λ are 9,0,9 in that order. So another answer to the question is

$$P^{-1}AP = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9 \end{bmatrix} \text{ where } P = \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -2 \\ 1 & -2 & 0 \end{bmatrix}$$

Multiples of u are also eigenvectors corresponding to λ_1 and combinations of v and w are also eigenvectors corresponding to λ_2. So I can also use 2u, 3v, v − 4w as cols of P and get

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} \text{ where } P = \begin{bmatrix} 4 & 0 & -4 \\ 2 & 6 & 10 \\ -4 & 3 & 1 \end{bmatrix}$$

etc.

example 1 continued

If possible, diagonalize A with a unitary P.

solution Possible since A is symmetric. Use the Gram Schmidt process on v,w to get

$$\vec{u_1} = v = (0,2,1)$$

$$\vec{u_2} = w - \frac{u_1 \cdot w}{u_1 \cdot u_1} u_1 = w - \frac{-4}{5} u_1 = (1, -\frac{2}{5}, \frac{4}{5})$$

Normalize $\vec{u}$, $\vec{u_1}$, $\vec{u_2}$ and use them as the cols for P. Then P is unitary (actually, orthogonal) and

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} \text{ where } P = \begin{bmatrix} 2/3 & 0 & 5/\sqrt{45} \\ 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4\sqrt{45} \end{bmatrix}$$

warning

1. If the question says "diagonalize the matrix M" then you must find Λ and find P and write $P^{-1}MP = \Lambda$.

2. It isn't necessary to get orthonormal eigenvectors unless you are specifically requested to diagonalize A using a unitary (or orthogonal) P. In that case don't forget to get eigenvectors that are both orthogonal *and* unit length.

3. In example 1, don't write $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$ when you mean $P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$.

Don't forget the P and P^{-1}.

example 2

Suppose M is 3×3 with eigenvalues 7,6,5 and corresponding eigenvectors

$$u = (2,1,3), \quad v = (1,4,2), \quad w = (4,1,6).$$

Then M is diagonalizable and

$$P^{-1}MP = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 5 \end{bmatrix} \text{ where } P = \begin{bmatrix} 2 & 1 & 4 \\ 1 & 4 & 1 \\ 3 & 2 & 6 \end{bmatrix}$$

There is no way to diagonalize M with a unitary (or orthogonal) P since the eigenvectors corresponding to different λ's did not turn out to be orthogonal (and cannot be re-chosen to be orthog).

powers and roots of a diagonalizable matrix

Suppose

$$A = P\Lambda P^{-1} \text{ where } \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Then

$$A^n = P\Lambda^n P^{-1} = P \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} P^{-1} \tag{4}$$

And

$$\text{a square root of } A = P \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix} P^{-1} \tag{5}$$

Similarly for an $n \times n$ matrix.

proof of (4)

$$A^n = (P\Lambda P^{-1})^n = P\Lambda \underbrace{P^{-1}\, P}_{\text{cancel}}\Lambda \underbrace{P^{-1}\, P}_{\text{cancel}}\Lambda \underbrace{P^{-1}\, P}_{\text{cancel}}\Lambda P^{-1} \;\ldots\; P\Lambda \underbrace{P^{-1}\, P}_{\text{cancel}}\Lambda P^{-1} = P\Lambda^n P^{-1}$$

And by the nature of matrix multiplication, the n-th power of a diagonal matrix is found by raising the diagonal entries to the n-th power. So

$$\Lambda^n = \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix}$$

example 3

Let

$$A = \begin{bmatrix} 0 & 3 \\ 2 & -1 \end{bmatrix}$$

One way to find A^{100} is to multiply A by itself 100 times. Another way is to first diagonalize A.

The eigenvalues are $-3,2$ with eigenvectors $(1,-1)$, $(3,2)$. So

$$A = P\Lambda P^{-1} \text{ where } P = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix} \text{ and } \Lambda = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$$

So

$$A^{100} = P\Lambda^{100} P^{-1}$$

$$= \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} (-3)^{100} & 0 \\ 0 & 2^{100} \end{bmatrix} \begin{bmatrix} 2/5 & -3/5 \\ 1/5 & 1/5 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2}{5}\cdot 3^{100} + \frac{3}{5}\cdot 2^{100} & -\frac{3}{5}\cdot 3^{100} + \frac{3}{5}\cdot 2^{100} \\ -\frac{2}{5}\cdot 3^{100} + \frac{2}{5}\cdot 2^{100} & \frac{3}{5}\cdot 3^{100} + \frac{2}{5}\cdot 2^{100} \end{bmatrix}$$

And one square root of A is

$$P \begin{bmatrix} i\sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix} P^{-1}$$

Another square root of A is

$$P \begin{bmatrix} i\sqrt{3} & 0 \\ 0 & -\sqrt{2} \end{bmatrix} P^{-1}$$

etc.

summary of some important ideas (you should know the answers to these questions)

question 1 What does it mean to say that a matrix A is diagonalizable.

answer 1 It means that there exists a matrix P such that $P^{-1}AP$ is diagonal.

question 2 What does it mean to diagonalize A with a unitary matrix.

answer 2 It means finding a unitary matrix P such that $P^{-1}AP$ is diagonal.

PROBLEMS FOR SECTION 8.4

1. Diagonalize M if possible and do it with a unitary P if possible.

(a) $\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$ (eigenvalues are 2,1,1)

2. Find a diagonal matrix D and an invertible matrix Q (if possible) so that $Q^{-1}AQ = D$. And make Q unitary if possible.

(a) A has 2-fold eigenvalue $\lambda_1 = 2$ with eigenvectors u = (1,0,0,0), v = (2,3,1,0) and 2-fold eigenvalue $\lambda_2 = 3$ with eigenvectors w = (0,0,0,2), x = (0,1,−3,1).

(b) A has 2-fold eigenvalue $\lambda_1 = 0$ with eigenvectors u = (1,0,0,0), v = (2,3,1,0) and 2-fold eigenvalue $\lambda_2 = 3$ with eigenvectors w = (1,1,0,0), x = (0,0,0,1).

3. Let A be 3 × 3. True or False.

(a) If A has 3 distinct eigenvalues then A can be diagonalized.

(b) If A has only 2 distinct eigenvalues then A can't be diagonalized.

4. Suppose A and B can be simultaneously diagonalized meaning that there is a matrix P so that $P^{-1}AP = \Lambda_1$ and $P^{-1}BP = \Lambda_2$ where Λ_1 and Λ_2 are diagonal. Show that A and B commute (i.e., AB = BA).

5. Suppose A is 3×3 with eigenvalues 2,3,4 and corresponding eigenvectors
$u = (1,2,3)$, $v = (5,6,7)$, $w = (9,10,11)$.
(a) Diagonalize A in several ways.
(b) Find $|A|$.

6. Let K be skew Hermitian.
(a) Show that iK is Hermitian
(b) Show that K can be diagonalized so that P is unitary and Λ is diagonal with pure imaginary entries.

7. It's possible for a lucky non-Herm to be diagonalizable with a unitary P.
But show that only Herms can be diagonalized with a unitary P *and* a real Λ.
In other words, show that if $P^{-1}AP = \Lambda$ where P is unitary and Λ is diagonal with real entries, then A is Herm.

8. Find A^{∞} if A is 2×2 with eigenvalues $\lambda_1 = .2$ and $\lambda_2 = 1$ and corresponding eigenvectors $u = (1,3)$, $v = (2,4)$.

REVIEW PROBLEMS FOR CHAPTER 8

1. Let $M = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Diagonalize M if possible.

2. Suppose A has eigenvalue λ with corresponding eigenvector u.
 (a) Is 2λ also an eigenvalue of A.
 (b) Is 2u also an eigenvector of A.

3. Suppose AB has eigenvalue $\lambda \neq 0$ with corresponding eigenvector u. Show that λ is also an eigenvalue of BA and find the corresponding eigenvector.

4. Suppose u is an eigenvector of A with corresponding eigenvalue 3, and v is an eigenvector of A with corresponding eigenvalue 0.
 (a) Show that u is in the range of A.
 (b) v is in what well-known subspace associated with A.

5. You want to make up an exam question so that the eigenvectors and eigenvalues come out to be "even". In particular you want M to be 2×2 with eigenvalues 3,5 and corresponding eigenvectors (2,3) and (1,0).
Make up such a matrix M.

6. Suppose the eigenvalues of a 3×3 matrix M are $2 \pm 3i$, 4.
Decide if M might, must, can't be (a) Hermitian (b) invertible (c) unitary

7. Let $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$
 (a) Diagonalize A.
 (b) Use part (a) to find A^8.

8. You already know that a noninvertible matrix always has eigenvalue 0 (from the invertible rule in §8.2).
 Confirm this again by thinking about the product of the eigenvalues.

9. Let A be $n \times n$.
 Use the characteristic polynomial to show if A has eigenvalue 3 then 8A has eigenvalue $8 \cdot 3 = 24$ (in general, multiplying a matrix by k will multiply its eigenvalues by k).

10. Suppose M is a $20 \approx 20$ matrix
 1 is a 12—fold eigenvalue with 12 ind corresponding eigenvectors
 -1 is an 8-fold eigenvalue with 8 ind corresponding eigenvectors
 Find M^{-1} (and make the answer as simple as possible).
 Suggestion: Think diagonalization.

CHAPTER 9 QUADRATIC FORMS

SECTION 9.1 THE MATRIX OF A QUADRATIC FORM

quadratic forms and their matrix notation

If $q = a_1 x^2 + a_2 y^2 + a_3 z^2 + a_4 xy + a_5 xz + a_6 yz$

then q is called a quadratic form (in variables x,y,z). There is a q value (a scalar) at every point. (To a physicist, q is probably the energy of a system with ingredients x,y,z.)

The matrix for q is

$$A = \begin{bmatrix} a_1 & \frac{1}{2}a_4 & \frac{1}{2}a_5 \\ \frac{1}{2}a_4 & a_2 & \frac{1}{2}a_6 \\ \frac{1}{2}a_5 & \frac{1}{2}a_6 & a_3 \end{bmatrix}$$

It's the symmetric matrix A with this connection to q:

(1) $$q = \begin{bmatrix} x & y & z \end{bmatrix} A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

or equivalently

(2) $$q = \vec{x}^{\,T} A \vec{x} \quad \text{where } \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

example 1

If $q = 3x_1^2 + 6x_2^2 - 7x_3^2 + \pi x_4^2 - \sqrt{2}\, x_1 x_2 + 8x_1 x_3 + \pi x_3 x_4 + 10x_2 x_4 - \frac{1}{2} x_1 x_4$

then the matrix for q is

$$A = \begin{bmatrix} 3 & -\sqrt{2}/2 & 4 & -1/4 \\ -\sqrt{2}/2 & 6 & 0 & 5 \\ 4 & 0 & -7 & \pi/2 \\ -1/4 & 5 & \pi/2 & \pi \end{bmatrix}$$

warning

In example 1, do *not* write q = A. The quadratic form does not *equal* a matrix (q is a scalar quantity, not a matrix). What q *does* equal is $\vec{x}^{\,T} A\vec{x}$ where $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_4 \end{bmatrix}$

example 2

If q has matrix $\begin{bmatrix} 2 & 3 & 5 \\ 3 & 7 & -2 \\ 5 & -2 & 1/2 \end{bmatrix}$ then $q = 2x^2 + 7y^2 + \frac{1}{2}z^2 + 6xy + 10xz - 4yz$.

basis changing rule for the matrix of a quadratic form

Suppose q is a quadratic form in variables x,y,z with (old) matrix A. Let

$$u = u_1 i + u_2 j + u_3 k$$
$$v = v_1 i + v_2 j + v_3 k$$
$$w = w_1 i + w_2 j + w_3 k$$

be a new basis for R^3. Let P be the usual basis-changing matrix:

$$P = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}$$

Then the new matrix for q in the new coord system with variables X,Y,Z and basis vectors u,v,w is given by

$$\text{new} = P^T \text{ old } P$$

example 3

Suppose

(3) $$q = x^2 + 6xy + y^2.$$

Let

$$u = i + 2j$$
$$v = -i + 2j$$

Find the matrix for q w.r.t. basis u,v and express q in terms of new coordinates X and Y.

solution *method 1* (using an algebraic substitution)

Let

$$P = \begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix}$$

The old coordinates x,y and new coordinates X,Y are related by

$$\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} X \\ Y \end{bmatrix}$$

so

$$x = X - Y$$
$$y = 2X + 2Y$$

Substitute for x and y in (3):

$$q = (X-Y)^2 + 6(X-Y)(2X+2Y) + (2X+2Y)^2.$$

Multiply out and collect terms to get q in terms of X and Y:

$$q = 17X^2 + 6XY - 7Y^2.$$

The matrix for q w.r.t. basis u,v is $\begin{bmatrix} 17 & 3 \\ 3 & -7 \end{bmatrix}$.

method 2 (using the basis changing rule for the matrix of a quadratic form)

The matrix for q (w.r.t. the standard coord system) is $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$.

By the basis changing rule above, the new matrix for q is

$$P^T AP = \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 17 & 3 \\ 3 & -7 \end{bmatrix}$$

Use the matrix to express q in terms of X and Y:

$$q = 17X^2 + 6XY - 7Y^2$$

proof of the basis changing rule

$$q = \begin{bmatrix} x & y & z \end{bmatrix} A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}^T A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= \left(P \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \right)^T A \; P \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad \text{by the basis changing rule for coords of a vector}$$

$$= \begin{bmatrix} X & Y & Z \end{bmatrix} P^T AP \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad \text{by T rule}$$

So P^TAP satisfies (1) but using new coordinates.
Furthermore P^TAP is symmetric because

$$(P^TAP)^T = P^TA^TP^{TT} = P^TAP \quad (\text{since } A^T = A)$$

So P^TAP is the matrix of q w.r.t. the new basis.

warning

1. The matrix of a quadratic form must be symmetric.
2. If A is symmetric then P^TAP is also symmetric.
   If yours isn't, check your arithmetic.

congruent matrices (versus similar matrices)

If A is symmetric and P is invertible (so that its cols are independent and can serve as basis vectors) then P^TAP and A are called *congruent*.

The matrix P^TAP represents the same quadratic form as A, but w.r.t. a new basis consisting of the cols of P.

Remember from Section 6.2 that $P^{-1} AP$ and A are called *similar* (whether or not A is symmetric).

The matrix $P^{-1}AP$ represents the same linear transformation as A but w.r.t. a new basis consisting of the cols of P.

If the matrix P is orthogonal, which happens when the new basis is orthonormal, then $P^T = P^{-1}$ and congruence is the same as similarity.

mathematical catechism

question 1 What is a quadratic form.
answer A quadratic form say in variables x_1, x_2, x_3, x_4 is an expression of the form

$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + a_4x_4^2$ + cross terms like $b_1x_1x_2$, $b_2x_1x_3$ etc.

question 2 What is the matrix of a quadratic form.
answer The matrix of a quadratic form in variables $x_1, \ldots, x_n$ is the symmetric

matrix A such that $q = [x_1 \ \dots \ x_n] \, A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

question 3 What are congruent matrices.
answer Matrices A and B are congruent if A is symmetric and there exists an invertible matrix P such that $B = P^T AP$ (this automatically makes B symmetric too).

question 4 What are similar matrices.
answer Matrices A and B are similar if there exists a matrix P such that $B = P^{-1}AP$.

PROBLEMS FOR SECTION 9.1

1. Let $q = x^2 + 3y^2 + 8z^2 - 3xy - 4yz$.

Find the matrix A for q and write q in terms of A using matrix notation.

2. Write out the quadratic form which has matrix $\begin{bmatrix} 2 & 3 & 4 \\ 3 & 6 & 7 \\ 4 & 7 & 9 \end{bmatrix}$.

3. Suppose $q = x_2^2 + 3x_4 x_5$. Find the matrix for q.

4. Find the quadratic form with matrix $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 0 & 0 \\ 3 & 0 & 0 & -3 \\ 4 & 0 & -3 & 6 \end{bmatrix}$.

5. Let

$$q = x^2 + 4xy + 3y^2$$
$$\vec{u} = 3\vec{i} + \vec{j}$$
$$\vec{v} = 2\vec{i} - \vec{j}.$$

(a) Find the new formula for q w.r.t. the basis u,v using the basis changing rule for a quadratic form.
(b) Find the new formula for q w.r.t. the basis u,v again using an algebraic substitution.
(c) Suppose a point has coords X=1, Y=2 w.r.t. basis u,v.
Find the value of q at the point two ways, using its X,Y coordinates and then again using its x,y coordinates.

6. Start with $q = x^2 + 3xy - 5y^2$ and make the change of variable

$$X = x - y$$
$$Y = x + y$$

(a) Find q in terms of X and Y just with algebra.
(b) What new basis is involved when you use variables X and Y.
(c) Find q in terms of X and Y again using the basis changing rule for q.
(d) Find the (old) matrix for q. What is the connection between q and the old matrix (write an equation beginning "q = ").
(e) Find the new matrix for q. What is the connection between q and the new matrix (write an equation beginning "q = ").
(f) What's the connection between the old matrix for q and the new matrix for q.

7. Start with $q = x^2 + 4xy - y^2$ and make the change of variable

$$x = 2X - Y$$
$$y = X + 3Y$$

(a) Find q in terms of X and Y just with algebra.
(b) What new basis is involved when you use coordinates X and Y.
(c) Find q in terms of X and Y again using the basis changing rule for q.

8. Suppose the X-axis is the same as the x-axis and the Y-axis is found by rotating the y-axis clockwise 45°. Use the basis changing rule for a quadratic form to find
(a) the new formula for $q = x^2 + y^2$.
(b) the old formula for $q = XY$.
(c) the equation of the circle $x^2 + y^2 = 1$ in the new system.

9. In the usual x,y coord system, q is $2x^2 + 3xy + 4y^2$.
Switch to a new X,Y coord system which has the same axes as before but new scales. If the old scale was the inch, on the new X–axis use the foot and on the new Y–axis use a half-inch. Find q in the new coord system.

SECTION 9.2 DIAGONALIZING A QUADRATIC FORM

diagonalizing q

Start with a quadratic form q, in say 3 variables, with matrix A.

Diagonalizing q means finding a new X,Y,Z coord system in which the formula for q has no cross terms, i.e., is of the form $aX^2 + bY^2 + cZ^2$.

Equivalently, diagonalizing q means finding an invertible matrix P so that the P^TAP, the new matrix for q, is diagonal.

first method for diagonalizing q: using eigenvalues

Suppose q is a quadratic form in variables x,y,z with matrix A. Since A is symmetric it has a complete set of *orthonormal* eigenvectors with corresponding real eigenvalues λ_1, λ_2, λ_3. Then w.r.t. a basis for R^3 of orthonormal eigenvectors of A,

$$q = \lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2$$

In other words, the new matrix for q w.r.t. the new basis is

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

proof

Let u,v,w be the orthonormal eigenvectors and let P be the matrix with cols u,v,w. Then

new matrix for q	$= P^T AP$	basis changing rule for q
	$= P^{-1}AP$	P is an orthog matrix since its cols are orthonormal
	$= \Lambda$	by (1) in Section 8.3

example 1

Let

(1) $$q = 5x^2 + 8y^2 + 5z^2 - 4xy + 8xz + 4yz$$

The matrix for q is

$$A = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}$$

It turns out that
$\lambda=0$ is an eigenvalue of A with corresponding eigenvector u = (2,1,−2);
$\lambda=9$ is a 2-fold eigenvalue with corresponding eigenvectors v = (0,2,1), w = (1,−2,0).
The Gram Schmidt process on v,w produces

$$v_1 = (0,2,1), \quad w_1 = (1, -\tfrac{2}{5}, \tfrac{4}{5})$$

So far, u, v_1, w_1 are orthog eigenvectors. Normalize them to get

$$r = (\tfrac{2}{3}, \tfrac{1}{3}, -\tfrac{2}{3}), \quad s = (0, \tfrac{2}{\sqrt{5}}, \tfrac{1}{\sqrt{5}}), \quad t = (\tfrac{1}{\sqrt{45}}, -\tfrac{2}{\sqrt{45}}, \tfrac{4}{\sqrt{45}})$$

In the new coord system with basis vectors r,s,t,

(2) $$q = 9Y^2 + 9Z^2.$$

The substitution that turns (1) into (2) is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \text{where } P = \begin{bmatrix} r & s & t \end{bmatrix}$$

$$x = \frac{2}{3} X + \frac{1}{\sqrt{45}} Z$$

$$y = \frac{1}{3} x + \frac{2}{\sqrt{5}} Y - \frac{2}{\sqrt{45}} Z$$

$$z = -\frac{2}{3} X + \frac{1}{\sqrt{5}} Y + \frac{4}{\sqrt{45}} Z$$

As a check, if you substitute for x,y,z in (1) you should get the new version of q in (2):

$$\begin{aligned} q = {} & 5\left(\frac{2}{3} X + \frac{1}{\sqrt{45}} Z\right)^2 + 8\left(\frac{1}{3} x + \frac{2}{\sqrt{45}} Y - \frac{2}{\sqrt{45}} Z\right)^2 + 5\left(-\frac{2}{3} X + \frac{1}{\sqrt{5}} Y + \frac{4}{\sqrt{45}} Z\right)^2 \\ & - 4\left(\frac{2}{3} X + \frac{1}{\sqrt{45}} Z\right)\left(\frac{1}{3} x + \frac{2}{\sqrt{45}} Y - \frac{2}{\sqrt{45}} Z\right) \\ & + 8\left(\frac{2}{3} X + \frac{1}{\sqrt{45}} Z\right)\left(-\frac{2}{3} X + \frac{1}{\sqrt{5}} Y + \frac{4}{\sqrt{45}} Z\right) \\ & + 4\left(\frac{1}{3} X + \frac{2}{\sqrt{45}} Y - \frac{2}{\sqrt{45}} Z\right)\left(-\frac{2}{3} X + \frac{1}{\sqrt{5}} Y + \frac{4}{\sqrt{45}} Z\right) \\ = {} & 9Y^2 + 9Z^2 \end{aligned}$$

warning

You need *orthonormal* eigenvectors to get q to be $\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2$. In example 1, if you use the original eigenvectors u,v,w (v and w are not orthog) as a basis, you won't get a diagonal q. And if you use eigenectors u, v_1, w_1 (orthogonal but not normalized) you will get a diagonal q but *not* $\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2$.

example 2

Diagonalize $q = 3x^2 + 3y^2 + 2xy$.
Find the new basis vectors and the change of variable that produces the diagonal q.

solution The matrix for q is

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

The eigenvalues are 2,4 with respective eigenvectors r = (1,−1), s = (1,1). They are orthog since A is symmetric. Normalize them to get

$$u = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \quad v = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

In the new coord system with basis u,v,

$$q = 2X^2 + 4Y^2.$$

The change of variable is

$$\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} X \\ Y \end{bmatrix} \quad \text{where } P = \begin{bmatrix} u & v \end{bmatrix}$$

$$x = \frac{1}{\sqrt{2}} X + \frac{1}{\sqrt{2}} Y$$

$$y = -\frac{1}{\sqrt{2}} X + \frac{1}{\sqrt{2}} Y$$

warning

Don't forget to use ortho*normal* eigenvectors if you claim that $q = \lambda_1 X^2 + \lambda_2 Y^2$.

Remember to normalize!

example 2 continued

Identify the graph of

$$3x^2 + 3y^2 + 2xy = 2$$

(temporarily obscured because of the cross term 2xy).

solution In the new X,Y cord system with basis u,v, the equation is

$$2X^2 + 4Y^2 = 2$$

so the graph is an ellipse (Fig 1).

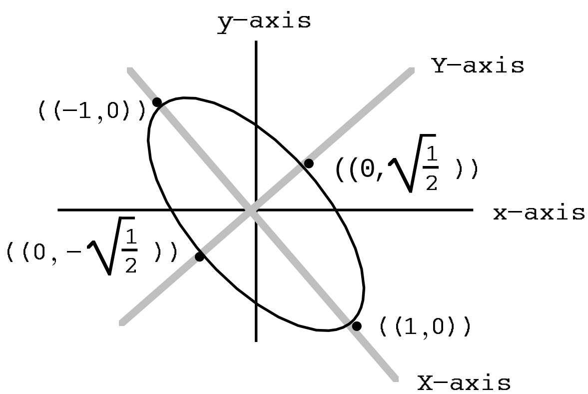


FIG 1

warning

Example 2 involved the quadratic form $q = 3x^2 + 3y^2 + 2xy$. The ellipse in Fig 1 is *not* the graph of q. It's the graph of the equation q = 2 (the ellipse is actually one of the *level sets* of q).

row/col operations on a matrix

A row/col operation is a pair of matching row and col ops such as

$$R_6 = 3R_1 + R_6$$
$$C_6 = 3C_1 + C_6$$

Note that row/col operations preserve symmetry; i.e., if you row/col operate on a symmetric matrix, the result is also symmetric.

row/col operation rule

Let A be the matrix of a quadratic form q w.r.t. the standard basis i,j,k. Do some row/col ops on A to get matrix B. Do just the col ops to I to get a new matrix P. Then B is the matrix of q w.r.t. a new basis consisting of the cols of P. In other words, $B = P^T AP$.

proof

Each *col* op corresponds to *right* multiplication by an elementary matrix (§1.3).
Each *row* op corresponds to *left* mult by an elem matrix.
When the operations are matching, the elem matrices turn out to be transposes of one another.
So row/col operations on A (say there are three of them) and the col ops on I amount to the scheme in Fig 2 where

I turns into $P = E_1E_2E_3$

A turns into $B = E_3^TE_2^TE_1^T AE_1E_2E_3 = (E_1E_2E_3)^T A (E_1E_2E_3) = P^TAP$

B is of the form P^TAP where P is invertible (elementary matrices are invertible and the product of invertibles is invertible), so B is the matrix for q in the new coord system whose basis vectors are the cols of P. QED

$$\begin{array}{ccc} A & \qquad\qquad & I \\ \Big\downarrow \text{row/col ops} & & \Big\downarrow \text{col ops} \\ B = E_3^TE_2^TE_1^T AE_1E_2E_3 & & P = IE_1E_2E_3 = E_1E_2E_3 \end{array}$$

FIG 2

second method for diagonalizing q: using row/col ops

> Suppose A is the matrix of a quadratic form q (say in variables x,y,z)
> Then you can do row/col ops to A to get a diagonal matrix
> $$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$
> Do just the col ops to I to get a new matrix P.
> Then with the cols of P as the new basis vectors,
> $$q = aX^2 + bY^2 + cZ^2$$

example 3

Let $q = 2x^2 + 5y^2 + 6z^2 + 8xy + 4xz + 2yz$.

Let A be the matrix for q. To diagonalize q using row/col ops, do the following row/col ops to A and do just the col ops to I.

	start with $A = \begin{bmatrix} 2 & 4 & 2 \\ 4 & 5 & 1 \\ 2 & 1 & 6 \end{bmatrix}$		start with $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$R_2 = -2R_1 + R_2$ $C_2 = -2C_1 + C_2$	$\begin{matrix} 2 & 4 & 2 \\ 0 & -3 & -3 \\ 2 & 1 & 6 \end{matrix}$	$\begin{matrix} 2 & 0 & 2 \\ 0 & -3 & -3 \\ 2 & -3 & 6 \end{matrix}$	$\begin{matrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$
$R_3 = R_1 + R3$ $C_3 = C_1 + C_3$	$\begin{matrix} 2 & 0 & 2 \\ 0 & -3 & -3 \\ 0 & -3 & 4 \end{matrix}$	$\begin{matrix} 2 & 0 & 0 \\ 0 & -3 & -3 \\ 0 & -3 & 4 \end{matrix}$	$\begin{matrix} 1 & -2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$
$R_3 = -R_2 + R_3$ $C_3 = -C_2 + C_3$	$\begin{matrix} 2 & 0 & 0 \\ 0 & -3 & -3 \\ 0 & 0 & 7 \end{matrix}$	$\begin{matrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{matrix}$	$\begin{matrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{matrix}$

After every step, you get a new symmetric matrix in the A column and a new set of basis col vectors in the I col. Each of the symmetric matrices represents q w.r.t. the new basis. For instance, after the first row/col op you know (but it isn't very useful) that

$$q = 2X^2 - 3Y^2 + 6Z^2 + 4XZ - 6YZ$$

in the new coord system with basis vectors (1,0,0), (−2,1,0), (0,0,1).

At the end you know (very useful) that

$$q = 2X^2 - 3Y^2 + 7Z^2$$

in the coord system with basis vectors

$$u = (1,0,0), \quad v = (-2,1,0), \quad w = (1,-1,1).$$

As a check, let $P = \begin{bmatrix} u & v & w \end{bmatrix}$. Then

$$P^TAP = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 2 \\ 4 & 5 & 1 \\ 2 & 1 & 6 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

so q does come out to be $2X^2 - 3Y^2 + 7Z^2$ in the coord system with basis u,v,w.

warning

In example 3, the diagonal coeffs 2,−3,7 are *not* eigenvalues of A and the cols of P are *not* eigenvectors. This was a *different* method of diagonalizing.

diagonalization process continued

You can continue the diagonalization in example 3 so that the coeffs are ±1's. Do some more row/col ops:

divide R_1 by $\sqrt{2}$; divide C_1 by $\sqrt{2}$

divide R_2 by $\sqrt{3}$; divide C_2 by $\sqrt{3}$

divide R_3 by $\sqrt{7}$; divide C_3 by $\sqrt{7}$

Do the col ops to P, where the earlier col ops left off. All in all, this turns A into

$$\begin{matrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

and turns I into

$$\begin{matrix} 1/\sqrt{2} & -2/\sqrt{3} & 1/\sqrt{7} \\ 0 & 1/\sqrt{3} & -1/\sqrt{7} \\ 0 & 0 & 1/\sqrt{7} \end{matrix}$$

So

$$q = X^2 - Y^2 + Z^2$$

in the coord system with basis vectors

$$\left(\frac{1}{\sqrt{2}}, 0, 0\right), \quad \left(-\frac{2}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0\right), \quad \left(\frac{1}{\sqrt{7}}, -\frac{1}{\sqrt{7}}, \frac{1}{\sqrt{7}}\right)$$

footnote This continuation in example 3 can also be accomplished algebraically by going from the latest X,Y,Z coord system to a new X_1, Y_1, Z_1 coord system by substituting $X = \frac{1}{\sqrt{2}} X_1$, $Y = \frac{1}{\sqrt{3}} Y_1$, $Z = \frac{1}{\sqrt{7}} Z_1$. You get

$$q = 2\left(\frac{1}{\sqrt{2}} X_1\right)^2 - 3\left(\frac{1}{\sqrt{3}} Y_1\right)^2 + 7\left(\frac{1}{\sqrt{7}} Z_1\right)^2 = X_1^2 - Y_1^2 + Z_1^2.$$

This substitution amounts to changing the scale on the X,Y,Z axes, i.e., to getting new basis vectors by multiplying the latest basis vectors by $1/\sqrt{2}$, $1/\sqrt{3}$ and $1/\sqrt{7}$ respectively (Section 2.4).

a generalization of the row/col operation rule for getting a new matrix for q

old rule

Start here with the matrix of q w.r.t.the usual basis i,j,k	Start with I, i.e., with [i j k]

Row/col operate on the left side. Col op on the right side.
You end up on the left side with a new matrix for q w.r.t. a new basis consisting of the cols on the right side.

more general rule

Start here with the matrix of q w.r.t. basis u,v,w	Start here with [u v w]

Row/col operate on the left side. Col op on the right side.
You end up on the left side with a new matrix for q w.r.t. a new basis consisting of the cols on the right side.

(This works whether or not you have reached a diagonal matrix on the left side although that's usually what you're aiming for.)

The more general version says you don't have to start at the "beginning", with the matrix for q w.r.t. the usual basis i,j,k.

example 4 (mixing two methods).

Suppose the matrix (2×2) for q has eigenvalues $\lambda = 2, -3$ with corresponding eigenvectors (2,1), (−1,2).

Diagonalize q so that the diagonal coeffs are ±1's and find the new basis.

solution

$q = 2X^2 - 3Y^2$ w.r.t. basis $u = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$, $v = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$.

Now pick up where the eigenvalue method left off and use row/col ops to make q diagonal with diagonal coeffs ±1's.

solution

matrix for q w.r.t. basis u,v	[u v]
$\begin{matrix} 2 & 0 \\ 0 & -3 \end{matrix}$	$\begin{matrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{matrix}$

Do these row /col ops on the left side and the col ops on the right side:

$R_1 = \frac{1}{2} R_1$, $C_1 = \frac{1}{\sqrt{2}} C_1$

$R_2 = \frac{1}{\sqrt{3}} R_2$, $C_2 = \frac{1}{\sqrt{3}} C_2$

You get

1	0	|	$\sqrt{2/5}$	$-1/\sqrt{5}$
0	−1	|	$1/\sqrt{10}$	$2/\sqrt{15}$

So $q = X^2 - Y^2$ w.r.t. basis $(\sqrt{2/5},\ 1/\sqrt{10})$, $(-1/\sqrt{15},\ 2/\sqrt{15})$.

graphing a conic in a NON-orthonormal coord system

In a non-orthonormal coord system, an ellipse can still be distinguished from a hyperbola in the usual way, by the signs of the coeffs in the equation.

But the usual way of distinguishing circles from non–circular ellipses doesn't work. For example, $2X^2 + 6Y^2 = 1$ can be a circle and $X^2 + Y^2 = 1$ can be a non-circular ellipse.

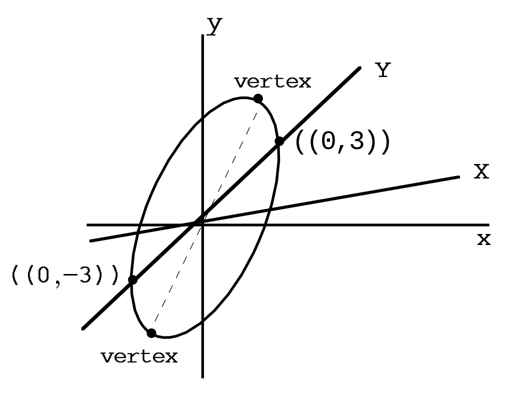
Furthermore, in a non-orthonormal system, the usual rules for finding vertices, foci, major and minor axes etc. no longer hold.

For example, in an orthonormal coord system, the graph of

$$\frac{X^2}{4} + \frac{Y^2}{9} = 1$$

is a (non-circular) ellipse with vertices $(0, \pm 3)$; the Y-axis is the major axis, the X-axis is the minor axis. But in a non-orthonormal system all you can conclude is that it is either a circle or an ellipse. Even if it is an ellipse, the points $(0, \pm 3)$ are not necessarily vertices, the X-axis and Y-axis are not necessarily the axes of the ellipse (Fig 2).

So to distinguish a hyperbola from an ellipse/circle, any diagonalization method is OK. But to graph accurately and locate vertices and axes of the conic and distinguish circles from non-circular ellipses, you have to use the eigenvector method to get q diagonal in a new *orthonormal* coord system.



possible graph of $\dfrac{X^2}{4} + \dfrac{Y^2}{9} = 1$

FIG 2

sign distribution law

Given a quadratic form q in n variables.

All diagonal versions of q have the same coefficient sign distribution, i.e., the same number of positive, negative and zero coeffs. This sign distribution is the same as the sign distribution of the eigenvalues of the matrix of q since one of the diagonal forms of q has the eigenvalues as the coeffs.

(proof omitted)

For the q in example 3 which turned out to be $2X^2 - 3Y^2 + 7Z^2$ in another coord system, every other diagonal form of q will have two positive and one negative coeff . In particular, if you find the eigenvalues of the matrix for q, two will be positive and one will be negative (but not necessarily 2,−3 and 7 specifically).

example 4

Suppose q is a function of four variables and in some coord system q is $2X_1^2 + 5X_3^2 - 7X_4^2$. Then every diagonalization of q has two positive coeffs, one negative coeff and one zero coeff. And the matrix for q has two positive eigenvalues, one negative eigenvalue (not necessarily 2,5 and −7 though) and one zero eigenvalue.

third method for diagonalizing q: completing the square

Look at the quadratic form

$$q = 3x^2 + y^2 + 4z^2 + 3xy + 6xz \tag{3}$$

To diagonalize q, first collect the x terms like this:

$$q = 3\Big[\;x^2 + (y+2z)x\;\Big] + y^2 + 4z^2$$

The coeff of x in the bracket is y+2z. Take half the coeff, square it and add it on to complete the square.

$$\begin{aligned} q &= 3\Big[\;x^2 + (y+2z)x + \left(\frac{y+2z}{2}\right)^2\;\Big] + y^2 + 4z^2 - 3\left(\frac{y+2z}{2}\right)^2 \\ &= 3\Big[x + \frac{y+2z}{2}\Big]^2 + \frac{1}{4}y^2 - 3yz + z^2 \end{aligned}$$

Now collect the y terms and complete the square again:

$$q = 3\Big[x + \frac{y+2z}{2}\Big]^2 + \frac{1}{4}\Big[y^2 - 12yz\Big] + z^2$$

The coeff of y in the second brackets is 12z. Take half the coeff, square it and add it on to complete the square.

$$\begin{aligned} q &= 3\Big[x + \frac{y+2z}{2}\Big]^2 + \frac{1}{4}\Big[\;y^2 - 12zy + 36z^2\Big] + z^2 - \frac{1}{4}\,36z^2 \\ &= 3\Big[x + \frac{y+2z}{2}\Big]^2 + \frac{1}{4}\Big[\;y - 6z\Big]^2 - 8z^2 \end{aligned}$$

Now let

$$X = x + \frac{y+2z}{2}, \quad Y = y - 6z, \quad Z = z. \tag{4}$$

Then

$$q = 3X^2 + \frac{1}{4}Y^2 - 8Z^2 \tag{5}$$

If you want the basis vectors corresponding to the new coords X,Y,Z, write (4) in matrix form:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1/2 & 1 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix}}_{P^{-1}} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Invert to get

(6)
$$P = \begin{bmatrix} 2 & -1/2 & -4 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}$$

The new basis vector are u = (2,0,0), v = (−1/2, 1, 0),(−4,6,1).

PROBLEMS FOR SECTION 9.2

1. (a) Let $q = x^2 + 4xy - 2y^2$. Use the eigenvalue method here.
   (i) Diagonalize q and find the basis for the new coord system in which q is diagonal.
   (ii) What change of variable produced the diagonal q.
   (iii) Find the old matrix for q and find the connection between q and the old matrix.
   (iv) Find the new matrix for q and find the connection between q and the new matrix.
   (v) What's the connection between the old matrix and the new matrix.
(b) Sketch the graph of $x^2 + 4xy - 2y^2 = 2$ and identify the major axis and vertices.

2. (a) Sketch the graph of $2x^2 + 2y^2 + 3z^2 + 4xy + 2xz + 2yz = 9$ by switching to a new orthonormal coord system.
The eigenvalues you need come out to be 2,5,0 with corresponding eigenvectors u = (−1,−1,2), v = (1,1,1), w = (1,−1,0).
(b) Why did you need eigenvalues in part (a). Wouldn't row/col ops or completing the square also have worked.
(c) What change of variables did you use in part (a). Express x,y,z in terms of X,Y,Z and vice versa.
(d) Repeat part (a) but change the 9 on the righthand side to −9.

3. Let $q = x^2 + 4xy - 2y^2$.
(a) Diagonalize q using row/col ops.
(b) Find the change of variable that produces the diagonal q in part (a) and check that it really works.
(c) Check that the new matrix for q really is P^T old P.

4. Let q have matrix $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 5 \\ 2 & 5 & 9 \end{bmatrix}$.
   (a) Diagonalize q using row/col ops and make the diagonal coeffs ±1's.
   (b) What change of variable produced the diagonal q in part (a).

5. Let $A = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 3 \\ 2 & 3 & 4 \end{bmatrix}$.
Without computing the eigenvalues, find their sign distribution (how many positives and how many negatives).

6. Suppose q has matrix A and the eigenvalues of A are 1/3 and −4 with corresponding eigenvectors u = (−4,3), v = (3,4).
Diagonalize q so that the coeffs are ±1's, and find the new basis vectors.

7. Diagonalize q by completing the square and find the new basis vectors.
   (a) $q = x^2 + 4xy - 2y^2$
   (b) $q = 2x^2 + 3xy + y^2$
   (c) $q = 3x^2 - 6y^2 + z^2 + 6xy + 18xz$
   (d) $q = 3x^2 + y^2 + 4z^2 + 2xy + 5xz + 6yz$

SECTION 9.3 DEFINITENESS

leading principal minors

The ℓpm's of a square matrix are the northwest subdeterminants.
For example, the ℓpm's of

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$$

are

$$|1|, \quad \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix}, \quad \begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{vmatrix}, \quad \begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{vmatrix}$$

positive definiteness

Let q be a quadratic form with matrix A (necessarily symmetric).
The following are equivalent (either all happen or none happen).
When they do happen then q and A are called positive definite.

(1) q is positive at every point (x,y,z) except at the origin $\vec{0}$ (where every q has value 0.)
(2) In every diagonal version of q the coeffs are positive.
In particular if you use row/col ops on A to get a diagonal matrix, the diagonal entries are positive.
(3) The eigenvalues of A are positive.
(4) The ℓpm's of A are positive.

positive semi (non-negative) definiteness

Let q be a quadratic form with matrix A (necessarily symmetric).
The following are equivalent (either all happen or none happen).
When they do happen then q and A are called non-negative definite or positive semi-definite.
(1) $q \ge 0$ at every point.
(2) In any diagonal version of q, the coeffs are ≥ 0.
In particular if you use row/col ops on A to get a diagonal matrix, the diagonal entries are ≥ 0.
(3) The eigenvalues of A are ≥ 0.
(4) The ℓpm's of A are ≥ 0.

The categories positive definite and positive semi-definite overlap: If A is positive definite then it is also positive semi-definite.

If in addition to (1)−(4) above, q actually *is* 0 somewhere besides the origin or in some diagonal version of q one of the coeffs actually *is* 0 or one of the eigenvalues of A actually *is* 0 or one of the ℓpm's actually *is* 0 then q and A are positive semi-definite but *not* positive definite.

negative definiteness

Let q be a quadratic form with matrix A (necessarily symmetric)
The following are equivalent (either all happen or none happen).
When they do happen then q and A are called negative definite.

(1) q is negative at every point (x,y,z) except at the origin $\vec{0}$ (where q = 0).
(2) In every diagonal version of q, the coeffs are negative.
In particular if you use row/col ops on A to get a diagonal matrix, the diagonal entries are negative.

(3) The eigenvalues of A are negative.
(4) The ℓpm's of A starting with the most northwesterly have the signs
$- + - + - + \ldots$

negative-semi (non-positive) definiteness

Let q be a quadratic form matrix with A (necessarily symmetric)
The following are equivalent (either all happen or none happen).
When they do happen then q and A are called non-positive definite or negative semi-definite.

(1) $q \le 0$ at every point.
(2) In any diagonal version of q, the coeffs are ≤ 0.
In particular if you use row/col ops on A to get a diagonal matrix, the diagonal entries are ≤ 0.

(3) The eigenvalues of A are ≤ 0.
(4) The ℓpm's of A starting with the most northwesterly are ≤ 0, ≥ 0, ≤ 0, ≥ 0, etc.

The categories negative definite and negative semi-definite overlap: If A is negative definite then it is also negative semi-definite.

If in addition to (1)−(4) above, q actually *is* 0 somewhere besides the origin or in some diagonal version of q one of the coeffs actually *is* 0 or one of the eigenvalues of A actually *is* 0 or one of the ℓpm's actually *is* 0 then q and A are negative semi-definite but *not* negative definite.

indefiniteness

If q and A are none of the above then they are called indefinite.

Only symmetric matrices are assigned a definiteness. So if a matrix is called positive, positive semi, neg, negative semi or indefinite you may assume that the matrix is symmetric to begin with.

For each kind of definiteness, the equivalence of (1)−(3) follows from the preceding sections.
The proof that (4) belongs on the list with (1)−(3) is too messy.

example 1

Let $q = x^2 - 3y^2 + 7xy$. Find the definiteness of q.

method 1 If $x = 0$, $y = 1$ then $q = -3$, negative. If $y=0$ and $x = 1$ then $q = 1$, positive. So by inspection, q is indefinite. (But it's not always this easy to find the definiteness by inspection.)

method 2 The matrix for q is

$$\begin{bmatrix} 1 & 7/2 \\ 7/2 & -3 \end{bmatrix}$$

The ℓpm's are $1, -61/4$ which makes q indefinite.

method 3 Start with the matrix for q and diagonalize it by adding $-\frac{7}{2}$ row1 to row2 and adding $-\frac{7}{2}$ col1 to col2 to get

$$\begin{bmatrix} 1 & 0 \\ 0 & -61/4 \end{bmatrix}$$

In another coord system (I don't care about the basis vectors), $q = X^2 - \frac{61}{4}Y^2$. There is one positive and one negative coeff. So q is indefinite.

method 4 The matrix for q has eigenvalues $\dfrac{-8 \pm 4\sqrt{65}}{8}$. One is positive and one is negative. So q is indefinite.

PROBLEMS FOR SECTION 9.3

1. Let

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Find the definiteness of A with several methods, for practice.

2. Find the definiteness. (a) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

3. True or False. If q is a quadratic form with all positive coeffs (e.g., $q = ax^2 + bxy + cy^2$ where a,b,c are positive) then q is positive definite.

4. Show that if A is negative definite then $-A$ is positive definite
(a) using (1) on the lists
(b) using (4) on the lists

5. Suppose A is positive definite. What about 3A.

6. Suppose q has matrix A, say 3×3. What can you conclude about the definiteness of A, the eigenvalues of A, and $|A|$ if
(a) in some new coord system, $q = -2X^2 - 3Y^2 - 5Z^2$
(b) in some new coord system, $q = X^2 - 2Y^2 + Z^2$

7. Let B be an arbitrary matrix.
(a) Show that B^TB is symmetric.
(b) Let q be the quadratic form with matrix B^TB. Write q in matrix notation.
(c) Show that B^TB is positive semi-definite by continuing from part (b) with matrix algebra until you can see that $q \ge 0$.
Suggestion: Use the fact that for column vectors u and v, $u \cdot v = u^Tv$ (Section 2.1).

8. Suppose A is symmetric. True or False.
(a) If A is positive definite then $|A|$ is positive
(b) If $|A|$ is positive then A is positive definite.

9. Let $q = 2x^2 + 3y^2 + bxy$. For what values of b will q be
(a) positive definite
(b) positive semi-definite
(c) positive semi-definite but not positive definite (i.e., *only* positive semi-definite)
(d) indefinite
(e) negative definite

10. (a) Show that if A is positive definite then A is invertible.
(b) Is the converse true?
(c) Suppose A is positive definite (so that by part (a), it is invertible). Show that the inverse must also be positive definite by thinking further about the eigenvalues of A^{-1}.

11. Let $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$. True or False.

(a) If A is positive definite then the diagonal entries a,d,f are positive.

(b) If the diagonal entries a,d,f are positive then A is positive definite.

SUMMARY OF BASIS CHANGING AND DIAGONALIZING

Read this before the 3rd exam and then again before the final exam.

basis changing for points, linear transformations, quadratic forms

Given a new coord system (i.e., given new basis vectors).

Let P be the basis changing matrix, with the new basis vectors as its cols.

Here's how to get the new coords of a point, the new matrix for a linear transformation T and the new matrix for a quadratic form q.

$$\begin{bmatrix} \text{new} \\ \text{coords} \\ \text{of} \\ \text{the} \\ \text{vector} \end{bmatrix} = P^{-1} \begin{bmatrix} \text{old} \\ \text{coords} \\ \text{of} \\ \text{a} \\ \text{vector} \end{bmatrix}$$

new Matrix for T = P^{-1} oldMatrix P (the old and new matrices are called *similar*)

new Matrix for q = P^T oldMatrix P (the old and new matrices are called *congruent*)

warning

The rule involves P^T when you get a new matrix for a quadratic form.

The rule involves P^{-1} when you get a new matrix for a transformation.

(If the new coord system is orthonormal then $P^T = P^{-1}$.)

diagonalizing a matrix

Diagonalizing M means finding a diagonal matrix Λ and an invertible matrix P so that $P^{-1}MP = \Lambda$.

Equivalently, diagonalizing M means finding a new basis so that the transformation which has matrix M w.r.t. the standard basis will have a diagonal matrix Λ w.r.t. the new basis.

There is only one method. The columns of P must be eigenvectors of M and the diagonal entries of Λ are the corresponding eigenvalues of M.

Not all matrices can be diagonalized (only the ones that have a complete set of eigenvectors).

Hermitian matrices are not only diagonalizable but you can do it with a real Λ and a unitary P provided you take the trouble to choose orthogonal normalized eigenvectors.

diagonalizing a quadratic form

Let q have matrix A.

Diagonalizing q means switching to a new coord system in which q has only square terms.

Equivalently, diagonalizing q means finding an invertible P so that P^TAP is diagonal.

Every q can be diagonalized (in many ways).

One diagonalizing method is to use as a basis (i.e., as the cols of P) a complete set of *orthonormal* eigenvectors of A in which case

$$q = \lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2.$$

This is the method to use if you want to graph an equation of degree 2, in two or three variables, accurately.

Other methods use row/col ops and completing the square and do not involve eigenvalues or eigenvectors or orthogonalizing or normalizing. These are the methods to use if all you care about is the definiteness of q (e.g., whether the graph of an equation of degree 2, in two variables, is an ellipse or hyperbola).

warning

To diagonalize q, you don't have to use eigenvectors at all. And the coeffs of the square terms in the new coord system don't have to be eigenvalues.

But if you want to turn q into $\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2$, you must not only use eigenvectors as the new basis then you must use orthonormal eigenvectors.

To diagonalize a matrix M you must use eigenvectors but they don't have to be orthogonal or unit length. It's just convenient if they are.

REVIEW PROBLEMS FOR CHAPTER 9

1. Let $q = 2x^2 - 4xy + 5y^2$

(a) Find the new formula for q in the coord system with basis $u = 2i$, $v = \frac{1}{3} j$ twice.

(i) Find the connection between X,Y and x,y and then use algebra
(ii) Use the basis changing rule

(b) Diagonalize q using three different methods.

(c) Sketch the graph of $2x^2 - 4xy + 5y^2 = 7$ and identify some of its significant features (in the old coord system).

2. Let

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

(a) Use two methods to show that A is positive definite.
(b) Find an invertible P so that $P^T AP = I$.

3. Suppose a quadratic form in variables x_1, x_2, x_3 has matrix A.

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Write the equation relating q,A, and $\vec{x}$.

4. Show that if A is congruent to B then $|AB| \geq 0$.

5. Let A be symmetric.
If row/col ops turn A into B and the col ops turn I into M, how are A,B,M related.

6. Suppose $q = x^2 - y^2 + 3xy$ in the standard coord system, and $q = -2X^2 + 3Y^2$ in another coord system with basis vectors u,v. Find the old matrix for q, the new matrix for q and write an equation relating the two matrices.

CHAPTER 10 FUNCTION SPACE

SECTION 10.1 VECTOR SPACES AND INNER PRODUCTS

axioms for an abstract vector space

A vector space V is a set of things, usually denoted $\vec{u},\vec{v},\vec{w},\ldots$ which can be added and multiplied by scalars (scalars are real or complex numbers usually denoted a,b,c,...).

The set V must be closed under addition and scalar multiplication meaning that if u and v are in V and k is any scalar then u + v and ku are also in V.

And the addition and scalar multiplication must obey the following axioms.

1. $u + v = v + u$
2. $(u + v) + w = u + (v + w)$
3. $a(u + v) = au + av$
4. $a(bu) = (ab)u$ (for example, $2(3u) = 6u$)
5. $au + bu = (a+b)u$ (for example, $2u + 3u = 5u$)
6. $1u = u$
7. There exists a zero vector denoted $\vec{0}$ so that for any u, $u + \vec{0} = u$
8. $u + -u = \vec{0}$ where $-u$ means $(-1)u$

Furthermore, subtraction of vectors is defined by $v - u = v + -u$.

If the scalars are real (resp.complex) then the vector space is called real (resp. complex).

example 1

The set of polynomials of degree 5 or less (e.g., $3x^5 + 2x - 3$, $x^2 + 2x - 1$, $x^5 + x^4 - 2x^3 + x^2 - x + 2$ etc.) is a vector space because these polynomials can be nicely added and multiplied by scalars (i.e., satisfy axioms 1–8 above).

example 2

The set of all $n \times m$ matrices is a vector space because same-size matrices can be nicely added and multiplied by scalars.

the vector spaces R^n and C^n

Let R^n be the set of of real n-tuples and let C^n be the set of complex n-tuples with addition and scalar multiplication defined as usual:

$$(u_1,\ldots, u_n) + (v_1,\ldots, v_n) = (u_1 + v_1,\ldots, u_n + v_n)$$

$$a(u_1,\ldots, u_n) = (au_1,\ldots, au_n)$$

Then R^n and C^n are vector spaces (real and complex respectively) because each is closed under addition and scalar multiplication, and axioms 1–8 hold.

dot products (inner products) for a vector space

Suppose an operation denoted $u\cdot v$ or $< u, v >$ combines two vectors in the space to produce a *scalar*. If the operation satisfies the following axioms then the operation is called an inner product or dot product or scalar product.

axiom 1 $v\cdot u = \overline{u\cdot v}$ (the overbar means conjugate)

For a real vector space, this simplifies to $v\cdot u = u\cdot v$

axiom 2 Dot products can be non-real but the dot product of a vector with *itself* must be a non-negative real number and can be 0 only if $u = \vec{0}$.

In other words,

$$u\cdot u > 0 \text{ if } u \neq \vec{0}$$

$$u\cdot u = 0 \text{ if } u = \vec{0}$$

axiom 3 $u\cdot(av) = a(u\cdot v)$
axiom 4 $u\cdot(v + w) = u\cdot v + u\cdot w$

The following properties can be proved from these axioms.

property 5 $(au)\cdot v = \bar{a}(u\cdot v)$
For a real vector space this simplifies to $(au)\cdot v = a(u\cdot v)$

property 6 $(u + v)\cdot w = u\cdot w + v\cdot w$
$(u + v)\cdot(w + z) = u\cdot w + v\cdot w + u\cdot z + v\cdot z$

the standard inner products for R^n and C^n

The standard inner product for R^n is

$$(u_1,\ldots,u_n)\cdot(v_1,\ldots,v_n) = u_1v_1 + \ldots + u_nv_n$$

It *is* an inner product because dot axioms 1–4 are satisfied.

Similarly, the standard inner product in C^n is

$$(u_1,\ldots,u_n)\cdot(v_1,\ldots,v_n) = \overline{u_1}\, v_1 + \ldots\ \overline{u_n}\, v_n$$

example 3 (another dot product in R^2)

Define an operation on R^2 as follows:

If $\vec{x} = (x_1, x_2)$ and $\vec{y} = (y_1, y_2)$ then $x \bullet y = 2x_1y_1 + 3x_2y_2$.

For example if $u = (3,4)$ and $v = (2,5)$ then $u \bullet v = 12 + 60 = 72$.

This *is* an inner product because dot axioms 1–4 hold.

orthogonal vectors

If $u\cdot v = 0$ then u and v are called orthogonal w.r.t. that particular dot product.

It's possible for u and v to be orthogonal w.r.t. to one dot product and not orthogonal w.r.t. a second dot product.

For example, if $p = (3,4)$ and $q = (6,-3)$ then p and q are not orthogonal w.r.t. to the standard dot product in R^2 because $p\cdot q = 18-12 = 6 \neq 0$. But they are orthogonal w.r.t. to the dot product in example 3 because $p \bullet q = 0$.

mathematical catechism

question 1 What does it mean to say that V is a vector space.
answer It means that the members of V can be added and multiplied by scalars so that certain axioms (the ones listed at the beginning of the section) are satisfied.
(It isn't necessary to memorize these axioms).

question 2 What does it mean to have a dot product on a vector space.
answer 2 It means that there is a way to combine two vectors to get a scalar so that the 4 dot axioms listed above are satisfied (yes, you should know these axioms)

PROBLEMS FOR SECTION 10.1

1. Show that once you choose axioms $v\cdot u = \overline{u\cdot v}$ and $u\cdot(av) = a(u\cdot v)$ you are doomed to have the property $(au)\cdot v = \bar{a}(u\cdot v)$.

Suggestion: Begin with $(au)\cdot v = \overline{v\cdot au}$.

2. A student looked at $x \bullet y = 2x_1y_1 + 3x_2y_2$ which example 3 claimed is a legal dot

product. She said it can't be right because if $u = (1,3)$ and $v = (-6,2)$ then $u \bullet v = -12 + 6 = -6$. But it should be 0 because u and v are orthogonal. Who is right?

3. Here are three attempts to make up some legal dot products in R^2. Do they work? If not explain why not. If $\vec{u} = (u_1,u_2)$ and $v = (v_1,v_2)$ then

(a) $u \bullet v = u_1v_2 + u_2v_1$

(b) $u \bullet v = 2u + 3v$???????

(c) $u \bullet v = 2u_1v_1 + 3u_2v_2$

4. (a) Find the dimension of the vector space of all 2×3 matrices by finding the "natural" basis.

(b) Show that the set of symmetric 3×3 matrices is a subspace of the vector space of all 3×3 matrices and find the dimension of the subspace.

(c) Is the set of invertible 2×2 matrices a subspace of the vector space of all 2×2 matrices.

5. Define a new inner product, denoted by o, in R^2 as follows.
If $x = (x_1, x_2)$ and $y = (y_1,y_2)$ then

$$x \circ y = x_1 y_1 - x_2 y_1 - x_1 y_2 + 2x_2 y_2$$

(a) Show that this new operation really is a legal inner product

(b) Are $\vec{i}$ and $\vec{j}$ orthogonal w.r.t. to the new inner product.

(c) Let $u = (2,3)$ and $v = (4,1)$. Are u and v orthogonal.

6. Let V be the set of all polynomials of degree 4 or less ; e.g., V contains

$2x^4 + 3x - 7$
x^3
$2x - 4$
etc.

Then V is a vector space because the polynomials can be nicely added and multiplied by scalars.

(a) Let $p = 5x^4 + 2x^2 - x + 2$.
Find a "natural basis" for V and then find the coords of p w.r.t. to your basis. What's the dimension of V.

(b) Let

$$q_1 = 3x^4,\ q_2 = 2x^3,\ q_3 = (x-3)^2,\ q_4 = x + 5,\ q_5 = 6$$

They are a basis for V (because they are independent and there are 5 of them).
Use a basis changing matrix to find the coords of $p = 5x^4 + 2x^2 - x + 2$ w.r.t. the new basis.

(c) Why isn't the set of polynomials of degree 4 (not of degree 4 *or less* but just of degree 4) *not* a vector space.

SECTION 10.2 FUNCTION SPACE

function spaces

Look at the set of real-valued functions f(x) of a real variable x.

The functions can be added; e.g., the sum of x^2 and sin x is the function $x^2 + \sin x$. And the functions can be multiplied by real scalars; e.g., 2 times sin x is the function 2 sin x. It can be shown that axioms 1–8 for a vector space from the preceding section are satisfied so the set is a (real) vector space.

Similarly, look at the set of all *complex-valued* functions of a *real* variable x; e.g.,

$$x^2 + ix$$
$$3x$$
$$ix^3$$
$$\cos x + i \sin x$$

Again, axioms 1–8 for a vector space are satisfied (allowing complex scalars) so the set is a (complex) vector space.

the standard inner product in function space

The standard dot product on the interval [a,b] is defined like this:

$$(1) \qquad \boxed{\langle f, g \rangle = \int_a^b \overline{f(x)}\; g(x)\, dx}$$

Drop the conjugate bars if f is real-valued

The standard dot product depends on the underlying interval [a,b].

For example, if $f(x) = 3x^2 + ix$ and $g(x) = 2x$ then on interval [0,1],

$$\begin{aligned}\langle f, g \rangle &= \int_0^1 (3x^2 - ix)(2x)\, dx \\ &= \int_0^1 (6x^3 - 2ix^2)\, dx \\ &= \int_0^1 6x^3\, dx - 2i \int_0^1 x^2\, dx \quad \text{(treat i like a constant)} \\ &= \left(\frac{6x^4}{4} - \frac{2ix^3}{3} \right) \Bigg|_0^1 \\ &= \frac{3}{2} - \frac{2}{3} i\end{aligned}$$

On the other hand, on interval [−1,1],

$$\langle f, g \rangle = \int_{-1}^1 (3x^2 - ix)(2x)\, dx = \frac{4}{3} i.$$

How do you know what interval to use? Out in life, the physics or engineering problem will involve an interval to begin with. In the meantime, a math problem will specify which interval.

It can be checked that dot axioms 1–4 from the last section hold so that (1) is a legitimate inner product.

Unless specifically stated otherwise, we always use the standard dot product in function space.

warning

If u is an abstract vector, don't write $\overline{u}$; there is no such thing as conjugating a vector *in general*. But it is OK if u is in *function space*; if u is the function $x^2 + ix^3$

then $\bar{u}$ is the function $x^2 - ix^3$.

functions orthogonal on the interval [a,b]

Two functions f and g are orthogonal on the interval [a,b] (w.r.t. the standard dot product) if their dot product on [a,b] is 0, i.e., if

$$\int_a^b \overline{f(x)}\; g(x)\, dx = 0 \qquad \text{(drop the conjugate bars if f is real-valued)}$$

Since the dot product depends on the interval, it is possible for two functions to be orthogonal on one interval and not orthogonal on another interval.

PROBLEMS FOR SECTION 10.2

1. Let $f(x) = 2x$ and $g(x) = x^2$.
Find $\langle f, g \rangle$ on interval [1,2].

2. If $f(x) = 2ix$ and $g(x) = 3x + ix$ find $\langle f, g \rangle$ and $\langle g, f \rangle$
(a) on interval [0,1] (b) on interval [−1,1]

3. Show that 1 and $\cos \frac{7\pi x}{L}$ are orthog on the interval [0,L].

4. Let $f(x) = \cos x + i \sin x$, $g(x) = 2 + 3i$. Are f and g orthogonal on
(a) interval $[0,2\pi]$
(b) interval $[0,\pi]$

5. Suppose f is a real-valued function and $f(a) = f(b)$.
Show that f and its derivative f' are orthogonal on [a,b].

6. (a) The standard inner product on [a,b] in (1) is the most popular one for the function space. But there are many others.
Define the following operation for real-valued functions:

$$(*) \qquad f \bullet g = \int_a^b f(x)\, g(x)\, x^2\, dx$$

The function x^2 in (*) is called a weight function.
Show that dot axioms 1 (real version) and 2 from the preceding section hold (and take my word for axioms 3 and 4).

(b) But suppose the weight function in (*) is changed from x^2 to x^3. Show that (*) would no longer be a legal dot product on interval [−1,0].

7. The natural basis for the vector space of polynomials of degree 4 or less is

$$x^4,\ x^3,\ x^2,\ x,\ 1$$

Show that the natural basis is not orthogonal on the interval [0,1] and then find a basis that *is* orthogonal on [0,1].

SECTION 10.3 NORMS

norm axioms

If there is an operation denoted $\|u\|$ which assigns to each vector u a *real* scalar satisfying the following axioms then the operation is called a norm.

axiom 1 $\|u\| > 0$ if $u \neq \vec{0}$

$\|u\| = 0$ if $u = \vec{0}$

axiom 2 $\|au\| = |a|\ \|u\|$

$|a|$ means the magnitude of the complex number a. It simplifies to the absolute value of a if a is real.

axiom 3 (triangle inequality) $\|u + v\| \le \|u\| + \|v\|$

warning

Norms are always real and furthermore are ≥ 0 (as opposed to dot products which can be non-real).

If you end up with a norm of -2 or $5i$ or $\sqrt{3i}$ you made a *mistake*.

the special norm associated with an inner product

Suppose a vector space has an inner product and you define the would-be norm

(1) $$\boxed{\|u\| = \sqrt{u \cdot u}}$$

It can be shown that this is a legal norm (i.e., satisfies norm axioms 1–3). It is called the norm associated with the inner product.

A vector space with an inner product can have norms in addition to the norm associated with its inner product.
But in this course unless specifically stated otherwise, the norm we use is always associated with a dot product.

proof that the would-be norm in (1) does satisfy norm axiom 1

First of all, you know that $u \cdot u \ge 0$ by dot axiom 1, so $\sqrt{u \cdot u}$ is the square root of a nonneg number so it is real, the first requirement for a norm.

Then norm axiom 1 holds because $\sqrt{u \cdot u}$ means the *non*-neg square root of the real number $u \cdot u$. And furthermore $\sqrt{u \cdot u} = 0$ iff $u \cdot u = 0$ which, by dot axiom 1, happens iff $u = \vec{0}$.

proof that the would-be norm in (1) does satisfy norm axiom 2

$\|au\| = \sqrt{au \cdot au}$ by definition of $\|\ \|$

$= \sqrt{\bar{a}a(u \cdot u)}$ by dot properties

$= \sqrt{\bar{a}a}\ \sqrt{u \cdot u}$ by the algebra of square roots

$= |a|\ \|u\|$ by the definition of $\|\ \|$ and the complex number rule $\bar{a}a = |a|^2$

proof that the would-be norm in (1) does satisfy norm axiom 3

Omitted. Much harder!

the norm associated with the standard dot product in R^n and C^n (the standard norm)

In R^n, the norm associated with the standard inner product is

$$\|u\| = \sqrt{u \cdot u} = \sqrt{u_1^2 + \dots + u_n^2}$$

In C^n, the norm associated with the standard inner product is

$$\|u\| = \sqrt{u \cdot u} = \sqrt{\overline{u_1}\, u_1 + \dots + \overline{u_n}\, u_n}$$

$$= \sqrt{|u_1|^2 + \dots + |u_n|^2}$$

the norm associated with the standard dot product in function space (the standard norm)

In the space of complex-valued functions with the standard inner product on interval [a,b], the associated norm is

(6) $$\boxed{\|f\| = \sqrt{< f, f >} = \sqrt{\int_a^b \overline{f(x)}\, f(x)\, dx} = \sqrt{\int_a^b |f(x)|^2\, dx}}$$

where $|f(x)|$ means the magnitude of the complex number $f(x)$.
If $f(x)$ is real-valued then (6) becomes

(6') $$\boxed{\|f\| = \sqrt{\int_a^b f^2(x)\, dx}}$$

The (standard) norm of a function depends on the interval. If you change the interval, the norm of a function changes.

example 1

On interval $[0,\pi]$,

$$\|\sin x\| = \sqrt{\int_0^\pi \sin^2 x\, dx} = \sqrt{\tfrac{1}{2}\pi}$$

On interval $[0,\pi]$, the normalized sine function with the awkward name $(\sin x)_{\text{unit}}$ is

$$\frac{\sin x}{\|\sin x\|} = \frac{\sin x}{\sqrt{\frac{1}{2}\pi}} = \sqrt{\frac{2}{\pi}} \sin x$$

example 2

If $f(x) = x^2 + 3i$ then

$$\|f\|^2 = \int_{x=0}^5 |f|^2\, dx = \int_{x=0}^5 (x^4 + 9)\, dx = 670$$

$$\|f\| = \sqrt{670}$$

$$f_{\text{unit}} = \frac{x^2 + 3i}{\sqrt{670}}$$

PROBLEMS FOR SECTION 10.3

1. Let $f(x) = x^2 + ix$. Find $\|f\|$ and the normalized f on
(a) interval [1,2] (b) interval [0,1]

2. A problem in the preceding section showed that 1 and $\cos\frac{7\pi x}{L}$ are orthog on the interval [0,L]. Normalize them to get an orthonormal pair on [0,L].
For reference: $\int \cos^2 u \, du = \frac{1}{2}(u + \sin u \cos u) = \frac{1}{2}u + \frac{1}{4}\sin 2u$

3. (a) Express in terms of integrals what it means for real-valued functions f(x) and g(x) to be orthonormal on interval [a,b].
(b) Repeat part (a) for complex-valued functions f(x) and g(x).

4. Find a,b,c so that the functions a and bx + c are orthonormal on [0,2].

5. Here's a dot product in R^2 (I checked that it is legal).

$$\text{If } \vec{u} = (u_1, u_2) \text{ and } v = (v_1, v_2) \text{ then } u \bullet v = 2u_1v_1 + 3u_2v_2.$$

Find the associated norm of u if u = (3,−2).

6. Here are some attempts to define a legal norm for R^n. Do they work? Explain
(a) If $\vec{u} = (u_1, \ldots, u_n)$ then $\|u\| = u_1 + \ldots + u_n$.
(b) If $\vec{u} = (u_1, \ldots, u_n)$ then $\|u\| = |u_1| + \ldots + |u_n|$.
(c) If $\vec{u} = (u_1, \ldots, u_n)$ then $\|u\|$ = biggest component − smallest component.
For example if u = (−3, −3, 8, 5) then $\|u\| = 8 - (-3) = 11$.

7. The collection of 2×2 matrices is a vector space. Give at least two reasons why $|M|$ (the determinant of M) doesn't count as a norm for the space even though the notation looks norm-like.

8. Define a new inner product, denoted by o, in R^2 as follows.
If $x = (x_1, x_2)$ and $y = (y_1, y_2)$ then

$$x \circ y = x_1y_1 - x_2y_1 - x_1y_2 + 2x_2y_2$$

(a) I checked out that dot axioms 1–4 are satisfied so this is a legal dot product. But for confirmation, check out axiom 2 yourself, the one that says

$$u \cdot u > 0 \text{ if } u \neq \vec{0}$$

$$u \cdot u = 0 \text{ if } u = \vec{0}$$

(b) Are $\vec{i}$ and $\vec{j}$ orthogonal w.r.t. to the new inner product.

(c) Let $u = \left(\frac{2}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right)$ and let $v = \left(\frac{4}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right)$

Show that u and v are orthonormal w.r.t. the new inner product and its associated norm denoted $\| \ \|_o$.

SECTION 10.4 ORTHOGONAL BASES FOR FUNCTION SPACE

orthogonal bases

Since the idea of orthogonality in function space depends on an interval, the idea of orthogonal basis for function space also depends on an interval.

The functions $y_1(x)$, $y_2(x)$, $y_3(x)$,... are an orthogonal basis on say interval [3,7] if the y's are orthogonal on the interval [3,7] and for any function f(x),

(1) $$f(x) = a_1y_1(x) + a_2y_2(x) + a_3y_3(x) + \ldots \quad \text{for } 3 \le x \le 7.$$

Note that we don't expect to be able to write f(x) as a combination of the y's for *all* x. Only for x in the interval on which the y's are orthogonal.

coordinates of a vector w.r.t. an orthogonal basis (same idea as in Section 2.4)

Suppose $y_1(x)$, $y_2(x)$, $y_3(x)$,... is an orthogonal basis on say interval [a,b].

Then

$$f(x) = a_1y_1(x) + a_2y_2(x) + a_3y_3(x) \text{ for } a \le x \le b.$$

where

(2) $$a_n = \frac{<y_n(x),\ f(x)>}{<y_n(x),\ y_n(x)>} = \frac{\int_a^b \overline{y_n}\ f(x)\ dx}{\int_a^b \overline{y_n(x)}\ y_n(x)\ dx}$$

(Drop the conjugate bars for real-valued functions.)

Furthermore, if the basis is ortho*normal* then the denominator in (2) is 1 and (2) simplifies to

(2a) $$a_n = <y_n(x),\ f(x)> = \int_a^b \overline{y_n}\ f(x)\ dx$$

example 1

It can be shown (coming up in Section 10.6) that the functions

(3) $$\sin\frac{\pi x}{L},\ \sin\frac{2\pi x}{L},\ \sin\frac{3\pi x}{L},\ \ldots$$

are an orthogonal basis for function space on the interval [0,L].

Let $f(x) = e^x$. Then for say $0 \le x \le 3$ we can write f(x) in terms of the basis vectors in (3) taking L=3.

In other words,

$$e^x = \sum_{n=1}^{\infty} c_n \sin\frac{n\pi x}{3} \quad \text{for } 0 \le x \le 3$$

where, by (2),

$$c_n = \frac{<\sin\frac{n\pi x}{3},\ e^x>}{<\sin\frac{n\pi x}{3},\ \sin\frac{n\pi x}{3}>} = \frac{\int_0^3 e^x \sin\frac{n\pi x}{3}\ dx}{\int_0^3 \sin^2\frac{n\pi x}{3}\ dx}$$

I started computing the coefficients using Mathematica (and rounded them off) and got

$$e^x = 7.02 \sin\frac{\pi x}{3} - 4.95 \sin\frac{2\pi x}{3} + 4.06 \sin\frac{3\pi x}{3} - 2.87 \sin\frac{4\pi x}{3}$$

(4)

$$+ 2.59 \sin\frac{5\pi x}{3} - 1.98 \sin\frac{6\pi x}{3} + 1.88 \sin\frac{7\pi x}{3} + \ldots \text{ for } 0 \le x \le 3$$

Fig 1 shows the graph of e^x together with the sum of the first 9 terms of the series in (4). For $0 \le x \le 3$, the sum is pretty close to x^2. The more terms you add, the more the sum looks like x^2 *but only for* $0 \le x \le 3$. (Actually it isn't good *at* x=0,3 either. The functions in (3) are not quite a basis. More about that in Section 10.6.)

```
numerator[n_]:= N[Integrate[E^x Sin[n Pi x/3], {x,0,3}]]
denominator[n_]:= Integrate[Sin[n Pi x/3]^2,{x,0,3}]
c[n_]:= numerator[n]/denominator[n]
sineSeries9 = Sum[c[n] Sin[n Pi x/3],{n,1,9}];
Plot[{E^x, sineSeries9},{x,-1,3.01}];
```

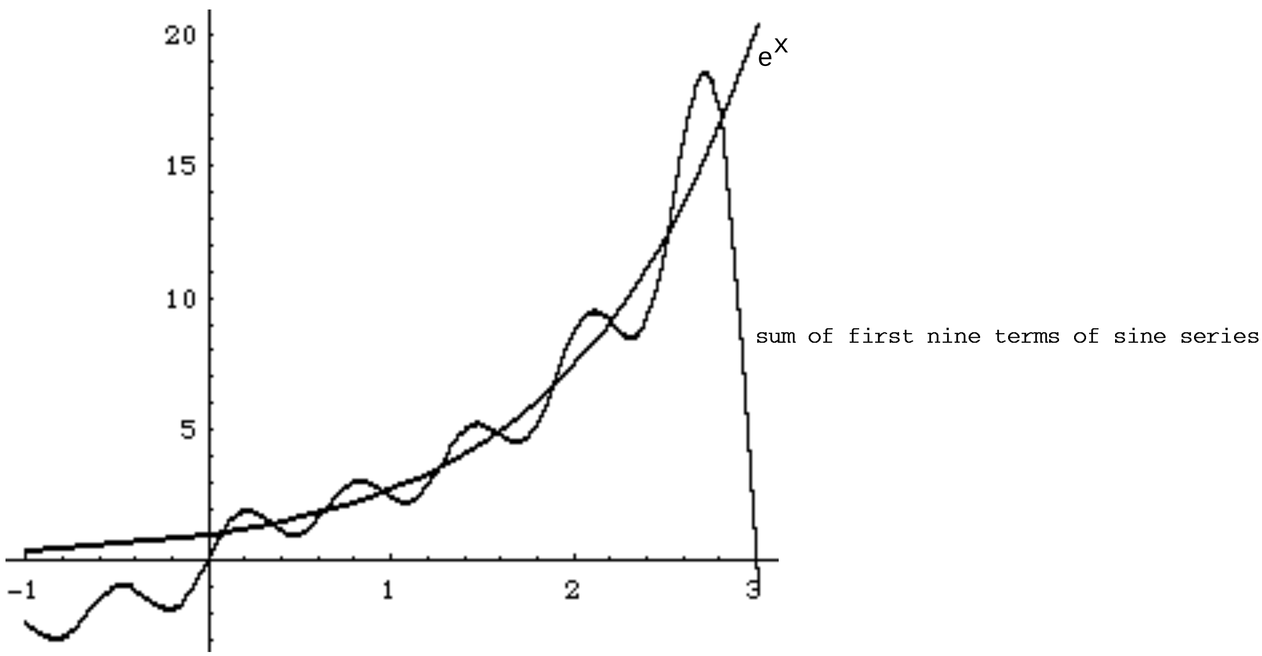


FIG 1

projections given an orthogonal basis for a subspace(same idea as in Section 3.2)

Suppose $y_1(x)$, $y_2(x)$, $y_3(x)$ are orthogonal functions on the interval [a,b].

Suppose f(x) is not a combination of y_1, y_2, y_3 but you would like the combination of y_1, y_2, y_3 that is closest to f(x). In other words, you would like to write

$$f(x) = (\text{approx})\ a_1y_1(x) + a_2y_2(x) + a_3y_3(x) \text{ for } a \le x \le b.$$

To get the best approximation, find a_1, a_2, a_3 so that $a_1y_1 + a_2y_2 + a_3y_3$ is the projection of f(x) into the subspace with orthogonal basis $y_1(x)$, $y_2(x)$, $y_3(x)$. In particular, choose

(5) $$a_i = \frac{< y_i(x),\ f(x) >}{< y_i(x),\ y_i(x) >} = \frac{\int_a^b \overline{y_i}\ f(x)\ dx}{\int_a^b \overline{y_i(x)}\ y_i(x)\ dx}$$

These values of a_1, a_2, a_3 make

$$\| a_1y_1(x) + a_2y_2(x) + a_3y_3(x) - f(x) \|$$

minimum; in particular, for real functions, these values of a_1, a_2, a_3 minimize the square root of

(6) $$\int_{x=a}^{b} \left(a_1y_1(x) + a_2y_2(x) + a_3y_3(x) - f(x) \right)^2 dx$$

and hence minimize the integral in (6) itself.

example 2
The functions 1 and cos x happen to be orthogonal on the interval $[0,\pi]$.
Find the best approximation to x^2 of the form $a + b \cos x$ for $0 \le x \le 3$.
solution

$$a = \frac{< 1,\ x^2 >}{< 1,\ 1>} = \frac{\int_0^\pi 1x^2\ dx}{\int_0^\pi 1^2\ dx} = \frac{1}{3}\pi^2$$

$$b = \frac{< \cos x,\ x^2 >}{< \cos x,\ \cos x >} = \frac{\int_0^\pi x^2 \cos x\ dx}{\int_0^\pi \cos^2 x\ dx} = \frac{-2\pi}{\pi/2} \quad \text{(I used Mathematica here)} = -4$$

So for $0 \le x \le 3$, the best approximation to x^2 of the form $a + b \cos x$ is
$\frac{1}{3}\pi^2 - 4 \cos x$ (Fig 2).

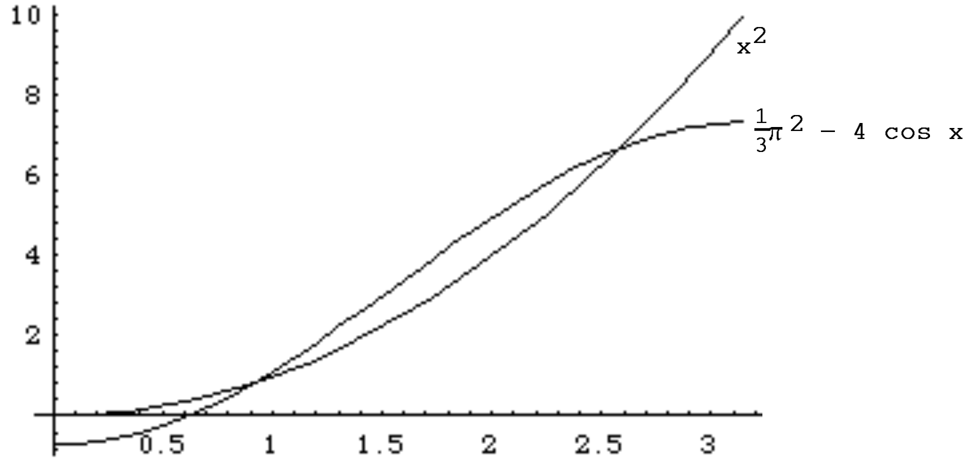


FIG 2

PROBLEMS FOR SECTION 10.4

1. It can be shown that the functions

$$y_0 = 1, \quad y_1 = \cos\frac{\pi x}{L}, \quad y_2 = \cos\frac{2\pi x}{L}, \quad y_3 = \cos\frac{3\pi x}{L}, \ \ldots$$

are an orthogonal basis on the interval [0,L].

(a) How would you check that y_5 and y_8 really are orthogonal on [0,L] (assuming you had a computer available to do computations).

(b) Find C_0 and C_n so that $e^x = C_0 + \sum_{n=1}^{\infty} C_n \cos\frac{n\pi x}{2}$ for $0 \le x \le 2$.

(But leave C_n in terms of integrals. Don't bother trying to do the computation.)

2. (a) Show that $3-4x$ and x^2 are orthogonal on the interval [0,1].

(b) Suppose you want to expand e^x in a tiny "series" of the form

$$e^x = a(3-4x) + bx^2 \quad \text{for } 0 \le x \le 1.$$

It can't be done because e^x isn't a polynomial. But find a and b so that your series gets as close as possible. (Leave the answer in integral form.)

(c) Express in terms of integrals what "as close as possible" means in part (b).

3. Go back to (4) where the function x^2 is written in terms of an orthogonal basis of sines for $0 \le x \le 3$.

Suppose you stop adding after 5 terms.

sum of first five terms

$$= 7.02 \sin\frac{\pi x}{3} - 4.95 \sin\frac{2\pi x}{3} + 4.06 \sin\frac{3\pi x}{3} - 2.87 \sin\frac{4\pi x}{3} + 2.59 \sin\frac{5\pi x}{3}$$

What is the official connection between x^2 and this sum.

SECTION 10.5 LINEAR OPERATORS

linear operators

Let T be a transformation (operator, function) mapping vectors to vectors. T is called linear if

(1) $T(u + v) = T(u) + T(v)$ and $T(au) = aT(u)$

or equivalently

(1a) $T(au + bc) = aT(u) + bT(v)$

linear operators on R^n and C^n

The linear operators on R^n and C^n are the matrix operators.

derivative of a complex-valued function of a real variable

To differentiate a complex-valued function of a real variable, just treat i like a constant. For instance

$$D_x(x^3 + ix^2) = 3x^2 + 2ix.$$

linear operators on function space

Suppose a (second order differential) operator T on a function space has the form

(2) $$\boxed{T(y) = ay'' + by' + cy}$$

where a, b, c are constants or functions of x. Then T is linear.

proof

$$\begin{aligned} T(y_1 + y_2) &= a(y_1 + y_2)'' + b(y_1 + y_2)' + c(y_1 + y_2) \\ &= a(y_1'' + y_2'') + b(y_1' + y_2') + c(y_1 + y_2) \\ &= \underbrace{ay_1'' + by_1' + cy_1}_{T(y_1)} + \underbrace{ay_2'' + by_2' + cy_2}_{T(y_2)} \end{aligned}$$

and

$$T(ky) = a(ky)'' + b(ky)' + c(ky) = aky'' + bky' + cky = k(ay'' + by' + cy) = kT(y) \text{ QED}$$

example 1

Suppose

$$T(y) = x^2 y'' + 3y' + e^x y.$$

For instance,

$$\begin{aligned} T(x^3 + ix^2) &= x^2(6x+2i) + 3(3x^2+2ix)+ e^x(x^3+ix^2) \\ &= e^x x^3 + 9x^2 + i(6x + e^x x^2) \end{aligned}$$

T has the form in (2) with $a = x^2$, $b = 3$, $c = e^x$, so T is linear.

eigenvalues and eigenvectors (same idea as in Section 8.1)

Let T be a linear operator on a vector space. Suppose u is a nonzero vector and λ is a scalar and

$$T(u) = \lambda u$$

Then λ is called an eigenvalue of T and u is the corresponding eigenvector.

eigenspaces (same idea as in Section 8 .1)

Let T be a linear operator on a vector space. Let k be a scalar.
Suppose u and v are eigenvectors corresponding to the eigenvalue λ.

Then u + v and ku are also eigenvectors corresponding to λ provided they are not $\vec{0}$ (i.e., provided $v \neq -u$ and $k \neq 0$).

The set of all eigenvectors corresponding to eigenvalue λ, together with the zero vector is a subspace called an eigenspace. And the axioms for a vector space continue to hold because they hold in the overall space. So an eigenspace is a vector space in its own right, a subspace of the overall space.

self-adjoint operators

Suppose a vector space has a dot product.
A linear operator T on the space is called self-adjoint (w.r.t. the dot product) if

(3) $\boxed{T(u)\cdot v = u\cdot T(v) \quad \text{for all vectors } u,v}$

self-adjoint operators on R^n and C^n

In §8.3, I showed that Hermitian matrices have the sliding property $Hu\cdot v = u\cdot Hv$. That means that Herm matrix operators (and symmetric matrix operators) are self-adjoint. It can also be shown that Herms are the *only* self-adjoint matrix operators.

In R^n, the symmetric matrices are the self-adjoint operators.
In C^n, the Hermitian matrices are the self-adjoint operators.

self-adjoint operators on function space

The definition of a self-adjoint operator involves a dot product.
And the standard dot product in function space involves an interval.
So in function space, the definition of a self-adjoint operator in (3) depends on an interval.

For instance with the standard dot product on interval [3,7] in function space, and with f(x) and g(x) in place of u and v, (3) becomes

(4) $$\int_3^7 \overline{Tf(x)}\; g(x)\, dx = \int_3^7 \overline{f(x)}\; Tg(x)\, dx \text{ for all functions f and g}$$

If (4) holds then we say that T is self-adjoint on the interval [3,7].
It's possible for a linear operator to be self-adjoint on one interval and not on another.

eigenvalues and eigenvectors of self-adjoint operators (same as in Section 8.3)

Suppose T is a self-adjoint operator on a vector space.

(a) The eigenvalues of T are real.
(b) Eigenvectors corresponding to different eigenvalues are orthogonal.
(c) For all practical purposes, there are enough eigenvectors to make an orthogonal basis for the space. *This is where all the famous orthogonal bases for function space come from.*

PROBLEMS FOR SECTION 10.5

1. Define an operator T on function space by T(y) = yy'. Show that T is not linear.

2. Let T be a linear operator on a vector space. Show that $T(\vec{0}) = \vec{0}$

Suggestion: Write $\vec{0}$ as $\vec{u} - \vec{u}$.

3. Suppose T is a linear operator on a vector space.
(a) Show that T preserves dependence meaning that if $u_1,...,u_n$ are dependent vectors then $T(u_1),..., T(u_n)$ are also dependent.
(b) Show that T does not necessarily preserve independence by producing a specific counterexample.

4. Define an operator T on function space by $T(y) = y''$. Show that $\sin\frac{7\pi x}{4}$ is an eigenfunction and find the corresponding eigenvalue.

5. Define two operators T and S on complex function space: Let
(a) $T(y) = xy$
(b) $S(y) = iy$.

For example, $T(\sin x) = x \sin x$ and $S(\sin x) = i \sin x$.
Each transformation is linear because it has the form in (2).
Show that T is self-adjoint on every interval [a,b], but S is not.

6. Suppose T is self-adjoint and
$$< T(u), v > = 8 - i, \quad < v, T(v) > = 7$$

Find whichever of the following you can with this information.
And state a reason for each significant step.
(a) $< v, T(u) >$
(b) $< u, T(v) >$
(c) $< u, v >$
(d) $< T(iu), v >$
(e) $< T(u+v), v >$
(f) $< T([2+3i]u), v >$
(g) $< v, T([2+3i]v) >$
(h) $< T(u+v), v >$

7. Let T be a self-adjoint operator on a complex vector space (meaning an abstract space, where the scalars are complex, but not necessarily C^n specifically). Show (in just a few steps) that $< u, T(u) >$ is real for all vectors u.
This is used in quantum mechanics where if T is say the energy operator, a self-adjoint operator, and u is the wave function of a particle then the inner product $< u, T(u) >$ represents the expected energy of the particle and you want it to come out real.

8. The set of all polynomials of degree 4 or less is a vector space (a subspace of function space) because the polynomials can be nicely added and multiplied by scalars. Define the operator T by

$T(p) = p'$ (T takes the derivative of a poly)

Then T has the form in (2) so T is linear.
(a) Is T self-adjoint on the interval [0,1].
(b) Find eigenvalues and eigenvectors of T.

9. The set of all 2×2 real matrices is a vector space because matrices can be nicely added and multiplied by (real) scalars.
Define an operator L on the space as follows: $L(A) = A^T$.
For example, if $B = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$ then $L(B) = \begin{bmatrix} 4 & 6 \\ 5 & 7 \end{bmatrix}$.
(a) Show that L is linear.
(b) Find the eigenvalues and eigenvectors (i.e., eigenmatrices) of L. In other words, look for nonzero matrices A such that $L(A) = \lambda A$. You should find two eigenvalues.
And find the dimension of each eigenspace.
(c) Define a dot product as follows (take my word for it that this is a legal dot

product):

If $A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ and $B = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ then $A\cdot B = a_1 b_1 + a_2 b_2 + a_3 b_3 + a_4 b_4$.

With respect to this dot product, is the operator L self-adjoint.

10. Let $T(y) = y \sin x$. Then T is linear because it has the form in (2), with $a = 0$, $b = 0$, $c = \sin x$. Show that T is self-adjoint on the interval $[a,b]$ (i.e., on any interval).

11. Look at the vector space of all 2×2 matrices.

Let $A = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$

Define operator T as follows.
If Q is a 2×2 matrix then $T(Q) = AQ$, i.e., T left multiplies by A.

For example, if $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ then $T(C) = AC = \begin{bmatrix} 3 & 6 \\ 12 & 16 \end{bmatrix}$

(a) Show (without doing a lot of algebra) that T is linear.

(b) Consider a dot product defined as follows:

If $C = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$ and $D = \begin{bmatrix} d_1 & d_2 \\ d_3 & d_4 \end{bmatrix}$

then $C\cdot D = c_1 d_1 + c_2 d_2 + c_3 d_3 + c_4 d_4$.

(It can be shown that this is a legal dot product.)

Is T self-adjoint w.r.t. this dot product.
If you say yes, prove it. If you say no, find a counterexample.

12. Define a linear operator on function space as follows:
$T(y) = x^2 y'' + xy'$.

Show that x^n is an eigenvector and find the corresponding eigenvalue.

10.6 THE SINE BASIS FOR FOR (A SUBSPACE OF) FUNCTION SPACE

a famous self-adjoint operator

Let $T(y) = y''$. Consider the interval $[0,L]$.
Look just at the subspace of functions $y(x)$ where $y(0) = 0$ and $y(L) = 0$.
Then T is a self-adjoint operator on $[0,L]$.

proof

I want to show that if f and g are in the subspace, i.e., if $f(0) = 0$, $f(L) = 0$, $g(0) = 0$, $g(L) = 0$ then

$$< T(f), g > = < f, T(g) >$$

To make it easier, I'll just do the case where $f(x)$ and $g(x)$ are real.

$$(1) \qquad < Tf, g > = \int_0^L f'' g \, dx$$

while

$$(2) \qquad < f, Tg > = \int_0^L f g'' \, dx$$

To show that the two integrals in (1) and (2) agree, use integration by parts.
For (1), let

$$u = g(x), \ dv = f''\,dx, \ du = g'(x)\,dx, \ v = f'$$

Then

$$(1') \qquad < Tf, g > = \int_0^L f'' g \, dx = \underbrace{f' g \Big|_0^L}_{\substack{0 \text{ because} \\ g(L)=0 \text{ and } g(0)=0}} - \int_0^L f' g' \, dx = - \int_0^L f' g' \, dx$$

Similarly, for (2), let

$$u = f(x), \ dv = g''\,dx, \ du = f'(x)\,dx, \ v = g'$$

Then

$$(2') \qquad < f, Tg > = \int_0^L f g'' \, dx = \underbrace{f g' \Big|_0^L}_{\substack{0 \text{ because} \\ f(L)=0 \text{ and } f(0)=0}} - \int_0^L f' g' \, dx = - \int_0^L f' g' \, dx$$

Look at (1') and (2') to see that $< Tf, g > = < f, Tg >$. QED

a famous orthogonal basis for function space

Let $T(y) = y''$.
Stick to the interval $[0,L]$ and to functions $y(x)$ satisfying $y(0) = 0$, $y(L) = 0$.

(a) The functions $\sin \frac{\pi x}{L}$, $\sin \frac{2\pi x}{L}$, $\sin \frac{3\pi x}{L}$, ... are eigenfunctions of T (corresponding to different eigenvalues).
(b) And furthermore they are the only eigenfunctions.

Since T is self-adjoint on the interval [0,L], the eigenfunctions in (a) are a basis for function space (actually just for the subspace of functions satisfying the conditions $y(0) = 0$, $y(L) = 0$). Furthermore they are orthogonal on the interval [0,L]. For every function f(x) we have

$$f(x) = a_1 \sin\frac{\pi x}{L} + a_2 \sin\frac{2\pi x}{L} + a_3 \sin\frac{3\pi x}{L} + \ldots \text{for } 0 \le x \le L$$

where

$$a_n = \frac{\left\langle \sin\frac{n\pi x}{L}, f(x) \right\rangle}{\left\langle \sin\frac{n\pi x}{L}, \sin\frac{n\pi x}{L} \right\rangle}$$

$$= \frac{\int_{x=0}^{L} f(x) \sin\frac{n\pi x}{L}\, dx}{\int_{x=0}^{L} \sin^2\frac{n\pi x}{L}\, dx}$$

footnote Actually the only functions f(x) that can be written as a sum of these sines for $0 \le x \le L$ are ones for which $f(0) = 0$ and $f(L) = 0$ because the sines are a basis only on the subspace of function space satisfying this condition.

In example 1 in Section 10.4, the sine series in (4) does not actually converge to x^2 for $0 \le x \le 3$ because x^2 is not 0 at x=3 (at least it is 0 at x=0). The series actually converges to x^2 only for $0 \le x < 3$. At x=3, the series converges to 0, not to x^2.

proof of (a)

I'll test $\sin\frac{3\pi x}{L}$ to see that it is an eigenvector. I'll find $T(\sin\frac{3\pi x}{L})$ and see if it comes out to be a multiple of $\sin\frac{3\pi x}{L}$:

$$T\left(\sin\frac{3\pi x}{L}\right) = \left(\sin\frac{3\pi x}{L}\right)'' = -\frac{9\pi^2}{L^2} \sin\frac{3\pi x}{L}$$

So $\sin\frac{3\pi x}{L}$ is an eigenfunction with corresponding eigenvalue $-\frac{9\pi^2}{L^2}$.

SOLUTIONS Section 1.1

1. $AB = \begin{bmatrix} -8 & 15 & 18 \\ -13 & 21 & 25 \end{bmatrix}$, $2A = \begin{bmatrix} 2 & 4 \\ 2 & 6 \end{bmatrix}$

BA and A + B don't exist.

2. AB, AC, BC, FQ, FA, AH, HA, BH, HQ, CF, QA

3. AB is 3×12
(a) If (AB)C is to exist then C must have 12 rows (doesn't matter how many cols)
(b) If C(AB) is to exist then C must have 3 cols (doesn't matter how many rows)

4. $AB = \begin{bmatrix} 13 & 18 & 24 \\ 9 & 17 & 42 \end{bmatrix}$. There is no BA.

5. $$\begin{bmatrix} 2 & 3 & 5 \\ 1 & -1 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$$

Also
$$x \begin{bmatrix} 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} 3 \\ -1 \end{bmatrix} + z \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$$

6. The sum is $a_{21} b_{15} + a_{22} b_{25} + a_{23} b_{35}$ which is like going across row 2 of A and down column 5 of B. It computes the entry in row 2, col 5 of AB.

7. $M\vec{i}$ is a column vector and in particular it's the first col of M.

8. (a) AB = [32], $BA = \begin{bmatrix} 10 & 15 & 20 \\ 12 & 18 & 24 \\ 2 & 3 & 4 \end{bmatrix}$

(b) Can't do AB; $BA = \begin{bmatrix} 10 & 15 \\ 12 & 18 \\ 2 & 3 \end{bmatrix}$

9. (a) 3×2

(b) Note that $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ is col 3 of the second factor. So $B \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} \text{col } 3 \\ \text{of} \\ \text{product} \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \\ 7 \end{bmatrix}$

10. Yes. For instance if A is 5×7 and B is 7×5 then AB and BA both exist. In general, AB and BA both exist if A is $p \times q$ and B is $q \times p$.

11. A^4B = AAAAB
= AAA(AB)
= AAA(5B)
= 5AAAB by property 4 of matrix multiplication
= $5AA(5B) = 5^2AAB = 5^2A(5B) = 5^3AB = 5^3(5B) = 5^4 B$

12. If $A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$ and $B = \begin{bmatrix} d & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & f \end{bmatrix}$

then AB and BA both equal $\begin{bmatrix} ad & 0 & o \\ 0 & be & 0 \\ 0 & 0 & cf \end{bmatrix}$.

13. Right multiplying matrix A by D has the effect of multiplying every entry in the first col of A by a, the second col of A by b, the third col of A by c; e.g., if

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{bmatrix} \text{ then } AD = \begin{bmatrix} 2a & 4b & 6c \\ 8a & 10b & 12c \\ 14a & 16b & 18c \end{bmatrix}$$

Left multiplying A by D has the effect of multiplying every entry in the first row of A by a, the second row of A by b, the third row of A by c.

14. $Ai = \begin{bmatrix} \text{col 1} \\ \text{of} \\ A \end{bmatrix}$, $Bi = \begin{bmatrix} \text{col 1} \\ \text{of} \\ B \end{bmatrix}$ by observation (3) about matrix mult.

Since Ai = Bi by hypothesis, col 1 of A must be the same as col 1 of B. Similarly, col 2 of A is the same as col 2 of B etc. So A = B.

15. Given that AB = BA.
To prove: $(6A - 4B)(2A + 3B) = (2A + 3B)(6A - 4B)$

proof: $(6A - 4B)(2A + 3B) = 12A^2 - 8BA + 18AB - 12B^2$
$= 12A^2 + 10AB - 12B^2$ since AB = BA

And
$(2A + 3B)(6A - 4B) = 12A^2 + 18BA - 8AB - 12B^2$
$= 12A^2 + 10AB - 12B^2$ since AB = BA

So
$(6A - 4B)(2A + 3B) = (2A + 3B)(6A - 4B)$ QED

warning about writing style (more of this coming up in Section 1.6)
Here is how *not* to write this proof:

bad style
$(6A - 4B)(2A + 3B) = (2A + 3B)(6A - 4B)$
$12A^2 - 8BA + 18AB - 12B^2 = 12A^2 + 18BA - 8AB - 12B^2$
$12A^2 + 10AB - 12B^2 = 12A^2 + 10AB - 12B^2$
TRUE
bad style

Any "proof" in mathematics that *begins* with what you want to prove as the first step and *ends* with something TRUE, is at best badly written and at worst incorrect.

What you should do instead to show that (6A−4B)(2A+3B) equals (2A+3B)(6A−4B) is work on one of them until it turns into the other or work on them both separately until they turn into the same thing (as I did in my answer).
The line (6A−4B)(2A+3B) = (2A+3B)(6A −4B) should be the *last* line of the proof, not the first.

16. Not necessarily. $(A + B)(A - B) = A^2 + BA - AB - B^2$ and doesn't equal $A^2 - B^2$ unless AB = BA which might happen but not necessarily.

17. (a) A(I + B) (b) (I + B)A

18. (a) The step from (A + I)(A − I) = 0 to A = I or A = −I is wrong because for matrices it is not true that if PQ = 0 then the only possibilities are P = 0 or Q = 0. Other possibilities may exist so there may be other solutions in addition to A = I and A = −I.
(b) Multiply each matrix by itself to see that you get I each time.

19. (a) I don't see how to factor. It isn't $B(A + C)$ for instance because this multiplies out to $BA + BC$ and not to $AB + BC$.

(b) $(A + C)B$

(c) $A(A + B)B$

(d) $A(A^2 + 2A + 6I)$ (Note The last term is 6I, not plain 6.)

(e) $(A + I)(A + 2I)$ Also $(A + 2I)(A + I)$

(f) Doesn't factor. It isn't $(A + 2B)(A + B)$ because this multiplies out to $A^2 + 2BA + AB + 2B^2$ which is not $A^2 + 3AB + 2B^2$.

20. $$\begin{bmatrix} \text{col 1} \\ \text{of} \\ C \end{bmatrix} = a \begin{bmatrix} \text{col 1} \\ \text{of} \\ A \end{bmatrix} + d \begin{bmatrix} \text{col 2} \\ \text{of} \\ A \end{bmatrix} + g \begin{bmatrix} \text{col 3} \\ \text{of} \\ A \end{bmatrix}$$

SOLUTIONS Section 1.2

1. (a) yes

(b) no. The third col is no good. It has to either be $\begin{matrix}0\\1\\0\\0\end{matrix}$ or of the form $\begin{matrix}a\\0\\0\\0\end{matrix}$

(c) no The second col is no good.
(d) yes
(e) yes

(f) no The last col is no good. It has to either be $\begin{matrix}0\\0\\1\end{matrix}$ or of the form $\begin{matrix}a\\b\\0\end{matrix}$

2. (a) Do $R_2 \leftrightarrow R_3$ and get $\begin{matrix}1&0&0&2\\0&1&0&0\\0&0&1&3\end{matrix}$

(b) $\begin{bmatrix}1&2&0&4&2\\0&0&0&0&3\\0&0&1&3&0\end{bmatrix} \sim \begin{matrix}1&2&0&4&2\\0&0&1&3&0\\0&0&0&0&3\end{matrix}$ $\quad R_2 \leftrightarrow R_3$

$\sim \begin{matrix}1&2&0&4&2\\0&0&1&3&0\\0&0&0&0&1\end{matrix}$ $\quad R_3 = \frac{1}{3} R_3$

$\sim \begin{matrix}1&2&0&4&0\\0&0&1&3&0\\0&0&0&0&1\end{matrix}$ $\quad R_1 = -2R_3 + R_1$

(c) $\begin{bmatrix}2&6&5\\4&10&0\\1&10&0\end{bmatrix} \sim \begin{matrix}1&3&5/2\\0&-2&-10\\0&7&-5/2\end{matrix}$ $\quad R_1 = \frac{1}{2} R_1,\ R_2 = -4R_1 + R_2,\ R_3 = -R_1 + R_3$

$\sim \begin{matrix}1&0&-25/2\\0&1&5\\0&0&-75/2\end{matrix}$ $\quad R_2 = -\frac{1}{2} R_2,\ R_1 = -3R_2 + R_1,\ R_3 = -7R_1 + R_3$

$\sim \begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}$ $\quad R_3 = \frac{2}{75} R_3,\ R_2 = -5R_3 + R_2,\ R_1 = \frac{25}{2} R_3 + R_1$

3. (a) Yes
(b) No. It's actually *two* row ops: $R_3 = 2R_3$ followed by $R_3 = R_4 + R_3$.
(c) No It's nothing at all.
(d) No It's *two* row ops: $R_3 = R_1 + R_3$ and $R_3 = R_2 + R_3$ (in either order).
(e) No It's *two* row ops: $R_2 = -R_2$ followed by $R_2 = 5R_4 + R_2$.
(f) No

4. $R_1 \leftrightarrow R_2,\ R_2 = \frac{1}{5} R_2,\ R_2 = -2R_1 + R_2$ (the opposite row ops in the opposite order)

To see that they work, do the original row ops to say $\begin{bmatrix}a&b\\c&d\end{bmatrix}$ and then do the new ones to see that the new row ops undo the original ones.

5. You can get A from I by doing the row op $R_1 = 2R_2 + R_1$ so A is an elem row matrix. (You can also get A from I by doing the col op $C_2 = 2C_1 + C_2$ so A also happens to be an elem col matrix.)
Each left multiplication by A does this row op. So

$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $A^2 = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$ (do the row op to A)

$A^3 = \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix}$ (do the row op to A^2) ,..., $A^{17} = \begin{bmatrix} 1 & 34 \\ 0 & 1 \end{bmatrix}$

6.(a) $E = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$, the elem matrix obtained from I by doing the row op $R_2 = 3R_1 + R_2$

(b) B can be obtained from A by the opposite row op, namely $R_2 = -3R_1 + R_2$.

Do this opposite row op to I and you get the elem matrix $F = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$.

7. (a) These row ops produced echelon form: $R_2 \leftrightarrow R_3$, $R_3 = \frac{1}{3} R_3$, $R_1 = -2R_3 + R_1$.
The corresponding elem matrices are

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The product $E_1E_2E_3$ does these row ops to A when it left multiplies A. So $B = E_3E_2E_1$
You can multiply the three E's together to get B. Or, as a short cut, you can start with E_1, do the second row op to E_1 (that produces E_2E_1) and then do the third row op (and you'll have $E_3E_2E_1$).

Either way, $B = \begin{bmatrix} 1 & -2/3 & 0 \\ 0 & 0 & 1 \\ 0 & 1/3 & 0 \end{bmatrix}$

(b) No, B is the product of elementary matrices but not an elementary matrix itself (because to get B from I, it takes *three* row ops, not one).

8. Call the product ABCD.
Matrices A,B and D are elem matrices.
D is right multiplying C. And D is the elem matrix obtained from I by adding 2 col 4 to col 1. So to get CD, do that col op to C.

footnote D is also the elem matrix obtained from I by adding 2 row 1 to row 4 but that point of view is not useful here since D is right multiplying C.

B is left multiplying CD. And B is the elem matrix obtained from I by switching rows 2 and 3. So to get BCD, do this row op to the matrix CD that you just got.
A is left multiplying BCD. And A is the elem matrix obtained from I by multiplying row 3 by 3. So to get ABCD, do this row op to the matrix BCD from the last step.
So all in all, to get ABCD, start with C and do these successive operations:

add 2 col 4 to col 1
switch rows 2 and 3
multiply row 3 by 3

Final answer is $\begin{bmatrix} 3 & 1 & 1 & 1 \\ 3x & x & x & x \\ 18 & 6 & 6 & 6 \\ 3y & y & y & y \end{bmatrix}$

9.(a) yes cols 1 and 4 have pivots
(b) no first col is no good
(c) no second col is no good
(d) no last col is no good
(e) yes cols 1,2,5 have pivots

SOLUTIONS Section 1.3

1. (a) The row numbers are 1 3,5,4,2. The diagrams shows that there are 4 inversions in the row numbers. So the product gets a plus sign.
(b) The row nubmers are 5 4 3 2 1. There are 10 inversions. Product gets a plus.

1 3 5 4 2

Problem 1

2 3 4 1

Problem 3

2. (a) $6 - -24 = 30$

(b) $7\begin{vmatrix} 0 & 3 \\ 1 & 17 \end{vmatrix} - 12\begin{vmatrix} 2 & 3 \\ 10 & 17 \end{vmatrix} + -4\begin{vmatrix} 2 & 0 \\ 10 & 1 \end{vmatrix} = -77$

(c) $-1\begin{vmatrix} 0 & 1 & 1 \\ 1 & -2 & 0 \\ 0 & 5 & 2 \end{vmatrix} + -3\begin{vmatrix} 2 & 1 & 1 \\ 1 & -2 & 0 \\ -1 & 5 & 2 \end{vmatrix} - 6\begin{vmatrix} 2 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 2 \end{vmatrix} + 1\begin{vmatrix} 2 & 0 & 1 \\ 1 & 1 & -2 \\ -1 & 0 & 5 \end{vmatrix} = \ldots = -1$

(d) -30 (product of diagonal entries)

3. (a) The row numbers corresponding to the product are 2 3 4 1.There are 3 inversions in the row numbers so the product gets a minus sign.
(b) Expanding down col 2 you get (among other terms)

$$-b\begin{vmatrix} 12 & 4 & d \\ a & -3 & 21 \\ 4 & c & 6 \end{vmatrix}$$

One term (of many) in the 3×3 det is (expand down col 1) $-a\begin{vmatrix} 4 & d \\ c & 6 \end{vmatrix}$.

One term in the 2×2 det is $-cd$
So overall, one term in the expansion is $-b\cdot-a:-cd = -abcd$
So the product abcd is in there, with a minus sign, which agrees with part (a).

4. $|A| = 7$, $|B| = 3$, $AB = \begin{bmatrix} 4 & 4 & -1 \\ 1 & 1 & 5 \\ 5 & 4 & 2 \end{bmatrix}$, $|AB| = 21$

5. There are zillions of counterexamples. For instance let
$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Then $A+B = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$ and $|A| = -2$, $|B| = 0$, $|A+B| = -2$ so for these matrices, $|A+B| \neq |A||B|$.

6. $(-1)^{27+29} = (-1)^{56} = +$

7. (a) $|AB| = |A||B| = -3$
(b) $6 + -\frac{1}{2} = 5\frac{1}{2}$
(c) Can't do with the given info
(d) $|3A| = 3^4|A| = 81\cdot 6 = 486$
(e) $3|A| = 3\cdot 6 = 18$
(f) $\left(\frac{1}{|A|}\right)^4 \cdot |A| = \left(\frac{1}{6}\right)^4 6 = \left(\frac{1}{6}\right)^3$

8. (a) $(-1)^4 \cdot -7 = -7$ (b) $(-1)^5 \cdot -7 = 7$

9. (a) No. Expand across row 1: det = $5\begin{vmatrix} 0 & 6 \\ 7 & 0 \end{vmatrix} = 5\cdot -6\cdot 7 = -5\cdot 6\cdot 7$

(b) Yes Expand across row 1

$$\det = -5\begin{vmatrix} 0 & 0 & 6 \\ 0 & 7 & 0 \\ 8 & 0 & 0 \end{vmatrix} = -5 \cdot -6\cdot 7\cdot 8 \text{ (as in part (a))} = 5\cdot 6\cdot 7\cdot 8$$

10. (a) The det of the zero matrix is the number zero so $|A^2| = 0$, $|AA| = 0$, $|A||A| = 0$. So $|A| = 0$ (the number zero).

(b) $|AA| = |I|$, $|A||A| = 1$, $|A|^2 = 1$, $|A| = \pm 1$

11. (a) The det of the echelon form is 0 since it has an all zero row. So the original det must be 0 also since row ops preserve zero-ness of the det

(b) The new matrix is triangular and its det is $2\cdot 4\cdot 7 = 56$. You can't conclude that the old det was 56 since you don't know what row ops were used and some row ops change the det. But you can conclude that the old det is nonzero because row ops preserve nonzero-ness

12. Do the row ops $R_1 = -2R_2 + R_1$, $R_2 = -R_3 + R_2$, $R_3 = -R_4 + R_3$ to get

$$\begin{bmatrix} 1-a & 0 & 0 & 0 \\ a-b & 1-a & 0 & 0 \\ b-c & a-b & 1-a & 0 \\ c & b & a & 1 \end{bmatrix}$$

The new det is $(1-a)^3$.

These are the type of row ops that don't change a det so $|A| = (1-a)^3$ too.

13. $$\begin{vmatrix} 1 & 3 & 6 \\ 2 & 4 & 4 \\ 1 & 2 & 8 \end{vmatrix} = 2\begin{vmatrix} 1 & 3 & 3 \\ 2 & 4 & 2 \\ 1 & 2 & 4 \end{vmatrix} \quad \text{(pull 2 out of col 3)}$$

$$= 4\begin{vmatrix} 1 & 3 & 3 \\ 1 & 2 & 1 \\ 1 & 2 & 4 \end{vmatrix} \quad \text{(pull 2 out of row 2)}$$

14. *Two* row ops were done: first $R_3 = 4R_3$ and then $R_3 = R_2 + R_3$

The second one doesn't change the det but the first one multiplies the det by 4.

15. (a) Expand by minors say down column 1:

$$\det = a\begin{vmatrix} a & b & 0 \\ 0 & a & b \\ 0 & 0 & a \end{vmatrix} - b\begin{vmatrix} b & 0 & 0 \\ a & b & 0 \\ 0 & a & b \end{vmatrix}$$

Each of the minor determinants is diagonal so is the product of the diagonal entries.

Answer is $a\cdot a^3 - b\cdot b^3 = a^4 - b^4$

(c) Similar to (a). It comes out to be $a^5 + b^5$.

16. (a) (i) $B = 3A$

(ii) $|B| = 3^4\ |A|$

(b) (i) Let E be the elementary matrix obtained from the 4×4 matrix I by tripling row 2:

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Then $B = EA$.

(ii) $|B| = 3|A|$

SOLUTIONS Section 1.4

1. $(A^TBC)^T = C^TB^TA^{TT} = C^TB^TA$

2. $|B| = |AA^T| = |A||A^T| = |A||A| = |A|^2$

3. (a) $A = \begin{bmatrix} & a_{1n} \\ & \\ a_{r1} & a_{rn} \end{bmatrix}$, $B = \begin{bmatrix} & b_{1n} \\ & \\ b_{q1} & b_{qn} \end{bmatrix}$

(b) To get the ij-th entry in BA^T, dot the i-th row in B (namely $b_{i1}\ b_{i2}\ \ldots\ b_{in}$) with the the j-th col of A^T (which is the j-th row of A, namely $a_{j1}\ a_{j2}\ \ldots\ a_{jn}$).

Answer is $b_{i1}a_{j1} + b_{12}a_{j2} + \ldots + b_{in}a_{jn}$.

SOLUTIONS Section 1.5

1. You can left multiply by A^{-1} on both sides to get

$$A^{-1}AB = A^{-1}AC$$
$$IB = IC$$
$$B = C$$

2. $P^{-1}PQR = P^{-1}C$ left multiply by P^{-1}

$QR = P^{-1}C$

$Q = P^{-1}CR^{-1}$ right multiply by R^{-1}

3. $EABF = I - CD$

Now to solve for B you need inverses for E,A,F. Assuming they exist,

$ABF = E^{-1}(I-CD)$

$BF = A^{-1}E^{-1}(I-CD)$

$B = A^{-1}E^{-1}(I-CD)F^{-1}$

4. All the answers are I because the product of a matrix and its inverse is always I

5. $(A^{-1}B^2A)^2 = A^{-1}B^2A \cdot A^{-1}B^2A = A^{-1}B^2\,B^2A$ since the AA^{-1} in the middle is I

$= A^{-1}$ BBBBA

$= A^{-1}$BBBAB since BA = AB

$= A^{-1}$BBABB since BA = AB

$= A^{-1}$BABBB again

$= A^{-1}$ABBBB again

$= B^4$ since $A^{-1}A = I$

6. (a) Use the quickie version for a 2×2

Inverse is $\frac{1}{4}\begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 3/4 & -1/2 \\ -1/4 & 1/2 \end{bmatrix}$

(b) No inverse. The det is 0 .

(c) det = 5, inv $= \frac{1}{5}\begin{bmatrix} -1 & 3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 5 \end{bmatrix} = \begin{bmatrix} -1/5 & 3/5 & 0 \\ 2/5 & -1/5 & 0 \\ 2/5 & -1/5 & -1 \end{bmatrix}$

(d) det = −4, inv $= \frac{1}{-4}\begin{bmatrix} -4 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1/2 \\ 0 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/2 \end{bmatrix}$

7. $\vec{x} = A^{-1}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1/5 & 3/5 & 0 \\ 2/5 & -1/5 & 0 \\ 2/5 & -1/5 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$

8. entry in row 3,col 2 of $A^{-1} = \dfrac{\text{cofactor of entry in row 2,col 3}}{\text{det M}}$

Use the cofactor of row 2, col 3 *not* row 3, col 2 because of the T in the formula

$$= \frac{-\begin{vmatrix} 1 & 1 & 0 \\ 0 & 0 & 2 \\ 6 & 8 & 12 \end{vmatrix}}{8} = \frac{1}{2}$$

9. $A^{-1} = [\frac{1}{4}]$ because $[4][\frac{1}{4}] = [1]$ which is the 1×1 version of I.

10. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$

11. Suppose AB = BA = I so that B is an inverse for A.
And suppose that AC = CA = I so that C is also an inverse for A.
Then

AB = I	by hypothesis
CAB = C	left multiply by C on both sides
IB = C	since CA = I by hypothesis
B = C	QED

12. Det $= (x-3)(x-1) - 8 = x^2 - 4x - 5 = (x-5)(x+1)$.
The det is nonzero for all x except 5 and -1
So the matrix is invertible if $x \neq 5$ and $x \neq -1$

13. Right multiply by A^{-1}:

$$B = \begin{bmatrix} 1 & 2 \\ 4 & 6 \\ 0 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 2 \\ 4 & 6 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 3 & -1 \end{bmatrix}$$

14. (a) Do these row ops:
$R_2 = -2R_1 + R_2$
$R_2 = -\frac{1}{5}R_2$
$R_1 = -3R_2 + R_1$
$R_3 = -R_2 + R_3$
$R_3 = -R_3$

They turn A into I. Do them to I and you'll get $A^{-1} = \begin{bmatrix} -1/5 & 3/5 & 0 \\ 2/5 & -1/5 & 0 \\ 2/5 & -1/5 & -1 \end{bmatrix}$

(b) Do these row ops
$R3 \leftrightarrow R_2$
$R_2 = \frac{1}{2} R_2$

$R_1 = -R_4 + R_1$

They turn A into I. Do them to I and you'll get $A^{-1} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

15. Do $R_1 = \frac{1}{2} R_1$, $R_2 = -R_1 + R_2$ and A turns into $\begin{matrix} 1 & 2 \\ 0 & 0 \end{matrix}$
This is echelon form and it isn't I so there is no inverse.

16. (a) A (b) A (c) $(A^3)^3 = AAA \cdot AAA \cdot AAA = A^9$

17. (a) $B^T A^T$
(b) $B^{-1} A^{-1}$
(c) ABABAB You can't simplify any further. You can't call it A^3B^3 because that would be switching the order of the factors which might change the product.

18. (a) $(A^{-1}B)^{-1} = B^{-1}(A^{-1})^{-1} = B^{-1}A$

(b) $\left[A(A+B)^{-1} \right]^{-1} = (A+B)A^{-1} = I + BA^{-1}$

19. $Q = (A^TA)^{-1} = A^{-1}(A^T)^{-1}$. Right multiply by A^T to get $QA^T = A^{-1}$

20. (a) $(A^T)^{-1} = (A^{-1})^T = \begin{bmatrix} 2 & 5 & 0 \\ 3 & 6 & 0 \\ 4 & 7 & 1 \end{bmatrix}$

(b) $(2A)^{-1} = \frac{1}{2}A^{-1} = \begin{bmatrix} 1 & 3/2 & 2 \\ 5/2 & 3 & 7/2 \\ 0 & 0 & 1/2 \end{bmatrix}$

(c) $|A| = \dfrac{1}{|A^{-1}|} = \dfrac{1}{-3}$ (to find det A^{-1}, expand across row 3)

(d) $|2A| = 2^3\,|A| = -8/3$

21. Let A be $n \times n$

(a) $|4A^{-1}| = 4^n\,|A^{-1}| = -4^n/3$

warning You have to preface this answer with "let A be $n \times n$".
The answer is meaningless unless you say what n represents.

(b) $(4A)^{-1} = \frac{1}{4}A^{-1}$ so $|(4A)^{-1}| = |\frac{1}{4}A^{-1}| = (\frac{1}{4})^n\,|A^{-1}| = (\frac{1}{4})^n \cdot -\frac{1}{3}$

22. *method 1* $|AB| \ne 0$ since AB is invertible.
But $|AB| = |A||B|$ so $|A||B| \ne 0$.
So $|A| \ne 0$ and $|B| \ne 0$.
So A and B are both invertible.

method 2

We are given that AB inverts.
So there is an $(AB)^{-1}$ such that $AB(AB)^{-1} = I$.
So $B(AB)^{-1}$ is a right-sided inverse for A which means that A is invertible (by the one-sided rule for testing for inverses in this section).

Similarly $(AB)^{-1}AB = I$ so $(AB)^{-1}A$ is a left inverse for B. That makes B invertible by the same shortcut.

warning It's wrong to start out like this: $(AB)^{-1} = B^{-1}A^{-1}$.
The rule does *not* say $(AB)^{-1} = B^{-1}A^{-1}$. The rule says that *if* A and B are invertible then AB is invertible and in that case $(AB)^{-1} = B^{-1}A^{-1}$. But in this problem you do not know that A and B are invertible. That's what you're trying to prove. So you have no right to quote $(AB)^{-1} = B^{-1}A^{-1}$.

23. Not invertible. Here's why.

$|AB| = |A||B|$ (det rule)
$= \text{who cares} \cdot 0$ ($|B| = 0$ since B is not invertible)
$= 0$

So AB is not invertible.

24. They are transposes of one another.
That's exactly what the rule $(A^{-1})^T = (A^T)^{-1}$ says.

25. It is not necessary to actually find A^{-1}. Since all you want is the determinant of A^{-1}, not A^{-1} itself, use the fact that $|A^{-1}| = 1/|A|$. The best way to find $|A|$ is to expand across row 3: $|A| = -8$. So $|A^{-1}| = -1/8$.

26. False. Here's a counterexample. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

Then A and B are each invertible (because they have nonzero dets). But A+B = 0 which is not invertible (its det is 0).

SOLUTIONS Section 1.6

1. (a) Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ Then $A^T = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$ and

$$A + A^T = \begin{bmatrix} \text{who cares} & d+b & g+c \\ d+b & \text{who cares} & f+h \\ g+c & f+h & \text{who cares} \end{bmatrix}$$ which is symm by inspection

(b) We want to show that $(A + A^T)^T = A + A^T$
Start with the left side and work on it.
$(A + A^T)^T = A^T + A^{TT} = A^T + A = A + A^T$ QED

2. (a) $A^T = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$

$$AA^T = \begin{bmatrix} \text{doesn't matter} & ad+be+cf & ag+bh+ci \\ ad+be+cf & \text{doesn't matter} & dg+eh+fi \\ ag+bh+ci & dg+eh+fi & \text{doesn't matter} \end{bmatrix}$$

Symmetric.
(b) I want to show that $(AA^T)^T = AA^T$. To do this, work on the left side until it turns into the right side.

$$\begin{aligned} (AA^T)^T &= A^{TT} A^T && \text{(T rule)} \\ &= AA^T && \text{(another T rule)} \quad \text{QED} \end{aligned}$$

3. I want to to show that $(B^TAB)^T = B^TAB$.
$(B^TAB)^T = B^TA^TB^{TT}$ (T rule) $= B^TA^TB$ (T rule) $= B^TAB$ (since A is symm)

4. (a) and (b) True, by inspection
(c) False. Here's a counterexample

Let $A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 5 \\ 5 & -2 \end{bmatrix}$

Then A and B are symm but $AB = \begin{bmatrix} \cdot & 4 \\ 23 & \cdot \end{bmatrix}$ which is not symm.

5. Let $A = \begin{bmatrix} a & e & f & g \\ e & b & h & i \\ f & h & c & j \\ g & i & j & d \end{bmatrix}$. Assume A inverts.

Then $A^{-1} = \dfrac{1}{|A|} \times$ adjoint of A.

For the purpose of showing that A^{-1} is symmetric, the diagonal entries in A^{-1} are irrelevant. I'll compute one pair of matching off-diagonal entries in A^{-1} to see if they are equal.

(*) row 1, col 2 entry in $A^{-1} = -\dfrac{1}{|A|} \begin{vmatrix} e & f & g \\ h & c & j \\ i & j & d \end{vmatrix}$

(Remember that there's a transpose in the formula for the adjoint so I got this determinant by crossing out row 2 and col 1 in A.)

And

$$(**) \quad \text{row 2, col 1 entry in } A^{-1} = -\frac{1}{|A|}\begin{vmatrix} e & h & i \\ f & c & j \\ g & j & d \end{vmatrix}$$

The arrays in (*) and (**) are transposes (this happened because the original A was symmetric). So the dets in (*) and (**) are equal. *You don't have to actually expand the two determinants to see that they are equal.* And in fact if you do expand out the two determinants you lose the idea that their being equal is not a coincidence and that the same argument would work for an $n \times n$ matrix.

And the attached signs are the same (both minuses), not a coincidence because the checkerboard pattern has the same sign in the ij-th spot as in the ji-th spot.

So these matching entries in A^{-1} are equal. Similarly for all the other matching pairs. So A^{-1} is symmetric. QED

6. I have to show that $AA^T = A^TA$.

I'll work on both sides until they agree with each other.

$AA^T = A(-A) = -A^2$

Also $A^TA = (-A)A = -A^2$.

So $AA^T = A^TA$.

7. Look at $(A^6)^T$:

$(A^6)^T = (A^T)^6$ (T rule)

$= (-A)^6$ (since A is skew symm)

$= (-A)(-A)(-A)(-A)(-A)(-A) = A^6$

So $(A^6)^T = A^6$. This makes A^6 symmetric, not skew symm.

SOLUTIONS Review problems for Chapter 1

1. D^9 is diagonal with diagonal entries $d_1^9, d_2^9, \ldots, d_{10}^9$. Just start multiplying D by itself and you'll see the pattern emerge.

2. $$\begin{bmatrix} 0 & & 0 & 1/6 \\ 0 & 0 & 1/5 & 0 \\ 0 & 1/4 & 0 & 0 \\ 1/3 & 0 & 0 & 0 \end{bmatrix}$$

(The fastest method is to just guess at a matrix which multiples the given matrix to give I .)

3. (a) If $A = \begin{bmatrix} 0 & -3 \\ 3 & 0 \end{bmatrix}$ then A is skew symm but $|A|$ is 9, not 0.

(b) Let $A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$ Find the det by expanding say across row 1.

$$|A| = -a \begin{vmatrix} -a & c \\ -b & 0 \end{vmatrix} + b \begin{vmatrix} -a & 0 \\ -b & -c \end{vmatrix} = -abc + abc = 0$$

(c) If A is skew symm then $A^T = -A$ so

$|A^T| = |-A|$

$|A| = (-1)^n\ |A|$ det rules

$|A| = -|A|$ since n is odd

$|A| = 0$ the only number that equals its opposite is 0

4. If $A^2 = 0$ (this means the zero matrix) then $|A^2| = |0|$, $|A||A| = 0$ (this means the number 0) so $|A| = 0$ so A is not invertible.

5. Not true. To disprove it you have to produce a counterexample.

If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ then $AB = \begin{bmatrix} 2 & \text{don't care} \\ \text{don't care} & 3 \end{bmatrix}$

trace AB = 5 but (trace A)(trace B) = $5 \cdot 0 = 0$.

6. (a) If $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ then $A \neq 0$, $B \neq 0$ but $AB = 0$.

(b) Given that $A \neq 0$, $B \neq 0$, A and B are invertible.
To show that $AB \neq 0$.

method 1 If A and B are invertible then AB is invertible (rule (3) of the algebra of inverses from §1.5). So AB can't be 0 (if it were the zero matrix it couldn't be invertible).

method 2

$|AB| = |A||B|$ (det rule)
= non-zero number·non-zero number (since A and B are invertible)
= non-zero (by the algebra of numbers)

So AB can't be the zero matrix (the zero matrix doesn't have a non-zero det).

7. If A is skew symm then $A^T = -A$ so $A + A^T = A + -A = 0$ (that's the zero *matrix*, not the number zero).

8. (a) Pull -1 out of col 2 . Det is -8

(b) This was gotten from the original by adding 2 row 1 to row 2 (which doesn't change the det) and tripling row 3 (which triples the det). New det is 24.

(c) This was gotten from the original by multiplying row 2 by 4, adding 2 row 1 to row 2 (no change in det) and tripling row 3. New det is $8 \cdot 4 \cdot 3 = 96$

9. Do these row ops.

$R2 = 3R1 + R2$
$R3 = -2R1 + R3$
$R3 = -R2 + R3$
$R2 = \frac{1}{3} R2$
$R1 = R2 + R1$

The echelon form is $\begin{matrix} 1 & 0 & 1/3 & 2/3 \\ 0 & 1 & -5/3 & -1/3 \\ 0 & 0 & 0 & 0 \end{matrix}$

10. (a) *argument 1* (indirect) Suppose A isn't invertible.
Then $|A| = 0$. And then $|AB| = |A||B| = 0\ |B| = 0$. So AB would not be invertible. But that's a contradiction since you are given that AB inverts.
So A must invert.

argument 2 (direct)

$|AB| = |A||B|$
$|AB| \neq 0$ since AB is invertible
So $|A||B| \neq 0$.
So $|A| \neq 0$ so A inverts.

warning It doesn't work to say A must invert because $(AB)^{-1} = B^{-1} A^{-1}$. You can't use the rule $(AB)^{-1} = B^{-1} A^{-1}$ until *after* you know that A and B invert.

(b) If you take $(AB)^{-1}$ which you know, and left multiply it by B which you know, you get $B(AB)^{-1}$ which is $B\,B^{-1}A^{-1}$ which is A^{-1}. So the answer is $A^{-1} = B(AB)^{-1}$.

11. (a) unreduced (to be reduced echelon form the last col must be 0 1 0)
(b) non-echelon (second col is no good)
(c) unreduced
(d) reduced
(e) non-echelon (last col is no good)
(f) unreduced

12. (a) $M^{-1} = \frac{1}{|M|}$ adj M so adj $M = |M|M^{-1}$.

So all I have to do is show that $|M|M^{-1}$ is invertible.
M^{-1} itself is invertible (its inverse is M).
And we know that $|M|$ is a *nonzero* number because M is invertible.
So $|M|M^{-1}$ is invertible because of the rule in Section 1.5 that if a matrix A is invertible and c is any non-zero number then cA is also invertible. QED

(b) (i) A inverts because $|A| \neq 0$ and in particular, $A^{-1} = \frac{1}{|A|}$ adj A. So

$$\text{adj } A = |A|\ A^{-1} = 3A^{-1}$$

So

$$|\text{adj } A| = |3A^{-1}| = 3^7\ |A^{-1}| = 3^7 \cdot \frac{1}{3} = 3^6$$

(ii) $|\text{adj } A| = |A|^{n-1}$

13. (a) I'll expand down col 1 to begin with and then keep expanding down col 1 in each subsequent minor det.

$$-1\begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{vmatrix} = -1\cdot -1\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -1\cdot -1\cdot -1\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -1\cdot -1\cdot -1\cdot 1 = -1$$

(b) There is only one product without a 0 factor, namely $1\cdot 1\cdot 1\cdot 1\cdot 1$. The corresponding row numbers are 2 5 3 1 4. Thereare 5 inversions so attached sign is minus. Final answer is -1.

14. (a) Yes because $|A^n| = |A|^n$.

(b) No. If the matrices say are 4×4 then her answer should be 3^4 times your answer. (If your answer was 0 then her answer is right. Lucky break.)

(c) Yes because $(A^T)^{-1} = (A^{-1})^T$.

(d) No. What she should do is take $\frac{1}{2}$ your answer since $(2A)^{-1} = \frac{1}{2} A^{-1}$.

15. $$|A| = 1\cdot\begin{vmatrix} x+1 & 1 & 2 \\ 1 & 1 & x+1 \\ 0 & 0 & 5 \end{vmatrix} = 1\cdot 5\cdot\begin{vmatrix} x+1 & x \\ 1 & 1 \end{vmatrix} = 1\cdot 5\cdot 1 = 5$$

$|A| \neq 0$ no matter what x is so A is invertible for all values of x.

(b) $|3A^{-1}| = 3^4\ |A^{-1}| = 81/|A| = 81/5$

16. (a) $R_1 = -2R_3 + R_1$

(b) $$E_1 = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad E_2 = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$E_1E_2 = I$ so E_1 and E_2 are inverses.

17. No. The det of the matrix is 0 (there are two identical rows; actually five identical rows). So it isn't invertible.

18. (a) The general rule does not say that $(CD)^{-1} = D^{-1}\ C^{-1}$.
The rule says that *if* C and D are invertible *then* $(CD)^{-1} = D^{-1}\ C^{-1}$.
In this case there is no chance that A^T and A are invertible since they aren't even square.

(b) $$Q^2 = Q\ Q = A\ \underbrace{(A^TA)^{-1}\ A^T\quad A}_{\text{look!}}\ (A^TA)^{-1}\ A^T$$

Look in the middle and you'll see $(A^TA)^{-1}\ A^T\ A$. This is of the form $(\text{thing})^{-1}$ thing. So it is I. So

$Q^2 = A\ I\ (A^TA)^{-1}\ A^T = = A(A^TA)^{-1}\ A^T$ *which is Q again.*
Conclusion: $Q^2 = Q$.

(c) From part (b) we have $Q^2 = Q$. So everywhere you see a Q^2 you can call it Q. And $Q^3 = Q^2\ Q$.
So

$Q^3 = Q\ Q = Q^2 = Q$.

So everywhere you see a Q^3 you can call it Q too. And

$Q^4 = Q^3Q = = QQ = Q^2 = Q$.

etc.

$Q^n = Q$ for all n. So $Q^{100} = Q$.

(d) Want to show that $Q^T = Q$.

$$Q^T = \left[A(A^TA)^{-1} A^T\right]^T$$

$$= (A^T)^T \left[(A^TA)^{-1}\right]^T A^T \quad \text{since } (ABC)^T = C^TB^TA^T$$

$$= A \left[(A^TA)^T\right]^{-1} A^T \quad \text{since } (A^T)^T = A \text{ and } (M^{-1})^T = (M^T)^{-1}$$

$$= A \left[A^TA\right]^{-1} A^T \quad \text{since } (CD)^T = D^TC^T$$

$$= Q$$

QED

SOLUTIONS Section 2.1

1. $u + v = (5,2,-2,11)$

$u \cdot v = 6 - 3 - 8 - 30 = -35$

$\|u\| = \sqrt{4 + 9 + 16 + 25} = \sqrt{54}$

$|u \cdot v| = 35$ (the absolute value of the number $u \cdot v$)

$$u_{unit} = \frac{u}{\|u\|} = \left(\frac{2}{\sqrt{54}}, \frac{3}{\sqrt{54}}, \frac{-4}{\sqrt{54}}, \frac{5}{\sqrt{54}} \right)$$

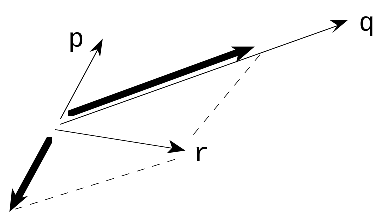
$$\frac{v \cdot u}{v \cdot v}\, v = \frac{-35}{50}\, v = \left(-\frac{105}{50}, \frac{35}{50}, -\frac{70}{50}, \frac{210}{50}\right)$$

2. (a) The old p has length $\sqrt{2}$ so the new p also has length $\sqrt{2}$. And the new p points like q. So

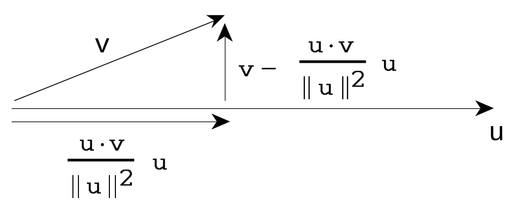
$$\text{new } p = \sqrt{2}\, q_{unit} = \sqrt{2}\left(\frac{2}{\sqrt{40}}, \frac{6}{\sqrt{40}}\right) = \left(\frac{2}{\sqrt{20}}, \frac{6}{\sqrt{20}}\right).$$

(b) vector projection $= \dfrac{p \cdot q}{q \cdot q}\, q = \dfrac{4}{\sqrt{40}} \left(-\dfrac{2}{\sqrt{40}}, -\dfrac{6}{\sqrt{40}}\right) = \left(-\dfrac{8}{40}, -\dfrac{24}{40}\right)$

3. Draw a parallelogram with diagonal r and sides parallel to p and q.
In the diagram below, r is the sum of the two thick arrows. It looks like r is about $-p + \frac{2}{3}q$.



Problem 3



Problem 4(a)

4. (a) The vector $\dfrac{v \cdot u}{\|u\|^2}\, u$ is the vector projection of arrow v onto arrow u.

The diagram shows the vector $v - \dfrac{v \cdot u}{\|u\|^2}\, u$ (use the triangle rule to subtract), and it's perpendicular to u.

(b) Dot the two vectors to see if you get 0.

Temporarily, let the scalar $\dfrac{u \cdot v}{\|u\|^2}$ be called k to make it easier to write. Then

$$\begin{aligned} u \cdot \left(v - \frac{u \cdot v}{\|u\|^2}\, u\right) &= u \cdot (v - ku) \\ &= u \cdot v - k(u \cdot u) && \text{dot property} \\ &= u \cdot v - k\|u\|^2 && \text{norm property} \\ &= u \cdot v - u \cdot v && \text{look at k and see that } k\|u\|^2 \text{ cancels down to } u \cdot v \\ &= 0 \end{aligned}$$

Since their dot product is 0, the vectors are orthogonal.

5. (a) $\|-2u\| = 2\|u\| = 6$

(b) $(u+v) \cdot (u-v) = u \cdot u - v \cdot v = \|u\|^2 - \|v\|^2 = 9 - 49 = -40$

(c) $\|u+v\| = \sqrt{(u+v) \cdot (u+v)}$ norm property

$= \sqrt{u \cdot u + 2\, u \cdot v + v \cdot v}$

$$= \sqrt{\|u\|^2 + 2\, u \cdot v + \|v\|^2}$$

$$= \sqrt{9 + 12 + 49} = \sqrt{70}$$

(d) $(u+3v)\cdot u = u\cdot u + (3v)\cdot u = \|u\|^2 + 3(v\cdot u) = \|u\|^2 + 3(u\cdot v) = 9 + 18 = 27$

6. Take the norm of $u/\|u\|$ and see if it's 1.

When you take the norm of $\frac{u}{\|u\|}$ use the rule $\|ku\| = |k|\,\|u\|$ (property (3) of norms) with $1/\|u\|$ playing the role of k; i.e., pull the scalar $1/\|u\|$ out from the norm signs. Since it's a positive scalar, I don't need the absolute value signs around it when it comes out

$$\left\| \frac{u}{\|u\|} \right\| = \frac{1}{\|u\|}\|u\| \quad \text{(careful use of property (3) of norms)}$$

$$= 1 \quad \text{(cancel the } \|u\|\text{'s)}.$$

7. On the left side, the multiplication between ku and v is the dot product (vector times vector). Similarly on the right side the multiplication of u and v is the dot product.

The multiplication of k and u on the left side is scalar mult (scalar times vector)

The multiplication of k and $u\cdot v$ on the right side is ordinary arithmetic multiplication, the product of two numbers.

8. $\|x + y\|^2 - \|x - y\|^2 = (x + y)\cdot(x + y) - (x - y)\cdot(x - y)$
$= x\cdot x + 2\,x\cdot y + y\cdot y - (x\cdot x - 2\,x\cdot y + y\cdot y)$
$= 4\,x\cdot y$ QED

warning *Don't write like this*

Don't write like this

$\|x + y\|^2 - \|x - y\|^2 = 4\,x\cdot y$

$(x + y)\cdot(x + y) - (x - y)\cdot(x - y) = 4\,x\cdot y$

$x\cdot x + 2\,x\cdot y + y\cdot y - (x\cdot x - 2\,x\cdot y + y\cdot y) = 4\,x\cdot y$

$4\,x\cdot y = 4\,x\cdot y$ TRUE

Don't write like this

Don't write like this

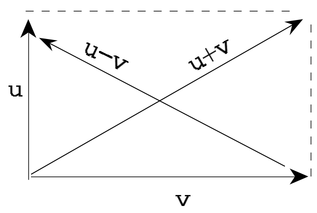
9. (a) The diagram shows u and v perpendicular (so that $u\cdot v = 0$) and shows u+v and u−v.The statement amounts to saying that in a rectangle, the two diagonals have the same lengths which we know is true.

(b) $\|u+v\| = \sqrt{(u+v)\cdot(u+v)}$ by the norm property $\|u\|^2 = u\cdot u$
$= \sqrt{u\cdot u + 2\,u\cdot v + v\cdot v}$ by dot properties
$= \sqrt{u\cdot u + v\cdot v}$ since $u\cdot v = 0$

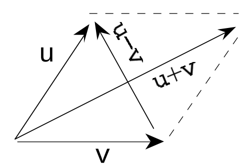
Similarly,

$\|u-v\| = \;= \sqrt{(u-v)\cdot(u-v)}$
$= \sqrt{u\cdot u - 2\,u\cdot v + v\cdot v}$
$= \sqrt{u\cdot u + v\cdot v}$

so $\|u+v\| = \|u-v\|$



Problem 9(a)



Problem 10(a)

10. (a) The diagram shows u and v with the same lengths (so that $\|u\| = \|v\|$). The statement amounts to saying that in a rhombus, the diagonal are perpendicular which we know is true.

(b) $(u+v)\cdot(u-v) = u\cdot u - v\cdot v = \|u\|^2 - \|v\|^2$ by norm property

$= 0$ since $\|u\| = \|v\|$ by hypothesis

11. Given $u\cdot v = 0$. Want to show that $u_{unit} \cdot v_{unit} = 0$

$$u_{unit} \cdot v_{unit} = \frac{u}{\|u\|} \cdot \frac{v}{\|v\|}$$

$$= \frac{1}{\|u\|}\ \frac{1}{\|v\|}\ (u\cdot v) \quad \text{by properties (2) and (3) of the dot product}$$

$$= \frac{1}{\|u\|}\ \frac{1}{\|v\|}\ (0) \quad \text{by hypothesis}$$

$$= 0$$

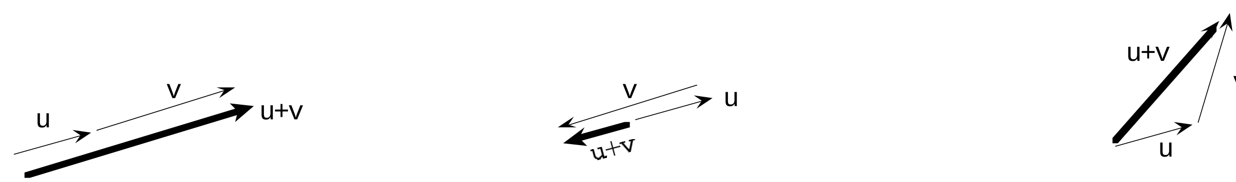
12. (a) Let $u = (1,0)$, $v = (0,1)$. Then $\|u\| + \|v\| = 1 + 1 = 2$

But $u + v = (1,1)$ and $\|u+v\| = \sqrt{2}$ so $\|u+v\| \neq \|u\| + \|v\|$.

So (*) can't be true in general in R^n.

(b) If v is a positive multiple of u then arrows u and v point the same way. The lefthand diagram shows the sum and you can see that the length of the sum is the u length plus the v length.

The other possibilities are that u and v point oppositely (center diagram) and they aren't parallel at all (righthand diagram). In these cases, $\|u+v\| < \|u\| + \|v\|$.



$\|u+v\| = \|u\| + \|v\|$ $\qquad$ $\|u+v\| < \|u\| + \|v\|$ $\qquad$ $\|u+v\| < \|u\| + \|v\|$

Problem 12(b)

(c) Let $v = ku$ where $k > 0$. I'll compute $\|u+v\|$ and $\|u\|+\|v\|$ and show that they are the same.

$\|u+v\| = \|u + ku\|$

$= \|(1+k)u\|$ property (5) of addition and scalar mult from §2.1

$= |1+k|\ \|u\|$ norm rule

$= (1+k)\|u\|$ don't need abs value signs since 1+k is positive

On the other hand,

$\|u\| + \|v\| = \|u\| + \|ku\|$

$= \|u\| + |k|\,\|u\|$ norm rule

$= \|u\| + k\|u\|$ don't need abs values since k is positive

$= (1+k)\|u\|$ ordinary factoring rule for numbers

So $\|\vec{u} + \vec{v}\| = \|\vec{u}\| + \|\vec{v}\|$.

SOLUTIONS Section 2.2

1. (a) True. Since u,v,w,x are ind then no one can be a comb of any of the others. In particular, u can't be a combination of v and w. Nor can v be a combination of u and w. Nor can w be a combination of u and v. So u,v,w are ind too.

(b) False. Let i = (1,0,0), j = (0,1,0), k = (0,01). Then i,j,k are ind but toss in u = (2,3,4) and the set of four vectors is dep because u is a comb of i,j,k.

(c) False. Take i = (1,0,0), j = (0,1,0), k = (0,01), u = (2,3,4) They are dep but take away u and the remaining set is independent.

(d) True. If u,v,w are dep then one is a comb of the others. So that same dependence relation ruins u,v,w,x.

(e) False. A counterexample is u = (1,0,0), v = (0,1,0), w = (2,3,0). Every pair of vectors is independent (i.e., if you look at any two vectors, neither is a multiple of the other) but the trio is dependent because w = 2u + 3v.

2. (a) There are 5 equations in 4 unknowns:

$$\begin{aligned} 2a + b + 5c - d &= 0 \\ 3a + 6c + 3d &= 0 \\ 4a + b + 9c + d &= 0 \\ 5a + 10c + 5d &= 0 \\ 6a + b + 13c + 3d &= 0 \end{aligned}$$

(b) The solution says

$$\begin{aligned} d &= \text{anything} \\ c &= \text{anything} \\ b &= -c + 3d \\ a &= -2c - d \end{aligned}$$

One possibility is the trivial solution d = 0, c = 0, b = 0, a = 0

But another solution (there are millions of others) is d = 1, c = 1, b = 2, a = −3

So $-3u + 2v + w + p = \vec{0}$.

So one of the many relations among the vectors is p = 3u − 2v − w.

Another solution is d=2, c=1, b=5, a=−4. So $-4u + 5v + w + 2p = \vec{0}$. So another relation among the vectors is $u = \frac{5}{4}v + \frac{1}{4}w + \frac{1}{2}p$.

3. The output would be a = 0, b = 0, c = 0, d = 0. Or as Mathematica actually writes it, you'd get this:

```
Out[1]
{{a -> 0, b -> 0, c -> 0, d -> 0}}
```

4. Look at the equation

(*) $$a(u + v) + b(v + w) + c(w + u) = \vec{0}$$

I want to show that a=0, b=0, c=0 is the only solution. Rearrange the equ to get

$$(a + c)u + (a + b)v + (b + c)w = \vec{0}$$

Since u,v,w are ind, by (3) we have

$$\begin{aligned} a + c &= 0 \\ a + b &= 0 \\ b + c &= 0 \end{aligned}$$

Subtract the first two equations and you get c − b = 0. Put this together with the third equation:

$$\begin{aligned} c - b &= 0 \\ b + c &= 0 \end{aligned}$$

This makes c = 0 and then a = 0 and b = 0.

The only solution to (*) turned out to be a = 0, b = 0, c = 0 .

So, by (3) again, u + v, v + w, w + u are ind.

5. (a) The system of equations $au + bc + cw + dp = \vec{0}$ is

$$\begin{aligned} a + b + c + d &= 0 \\ b + c + d &= 0 \\ c + d &= 0 \\ d &= 0 \end{aligned}$$

Start solving by looking at the last equation and working your way back up. The solution is d=0, c=0, b=0, a=0. So the vectors u,v,w,p are independent.

(b) u is not $\vec{0}$ and no vector after that is a combination of preceding vectors; for instance w isn't a combination of u and v since all combinations of u and v have a 0 in the 3rd spot but w has a 1 in the 3rd spot. So the vectors are ind.

6. *method 1* Look at the equation

(*) $$a\,Au + b\,Av + c\,Aw = \vec{0}.$$

I want to show that the only solution is a = 0, b = 0, c = 0.
Use matrix algebra on (*) to get

$$A(au + bv + cw) = \vec{0}$$

Then left multiply by A^{-1} on both sides to get

(**) $$au + bv + cw = \vec{0}$$

We know that u,v,w are ind. So by (3), applied to (**), we know that a=0, b=0, c=0.

So the only solution to (*) turned out to be a=0, b=0, c=0. So, by (3) again, Au, Av, Aw are ind.

method 2 Could one of Au, Av, Aw be a combination of the others? Suppose

$$Au = a\,Av + b\,Aw.$$

Then

$Au = A(av + bw)$ matrix algebra

$u = av + bw$ left multiply by A^{-1} on both sides

But this contradicts the hypothesis that u,v,w are independent. So you can't have Au = a Av + b Aw. Similarly no one of Au, Av, Aw can be a combination of the others. So they are ind.

7. *method 1* Try to solve $au + bv + cw = \vec{0}$.

If $au + bv + cw = \vec{0}$ then

$u\cdot(au + bv + cw) = u\cdot\vec{0}$ dot on both sides with u

$a(u\cdot u) + b(u\cdot v) + c(u\cdot w) = 0$ dot rules

$a(u\cdot u) = 0$ since $u\cdot v = 0$, $u\cdot w = 0$

$a = 0$ or $u\cdot u = 0$

But $u\cdot u \neq 0$ since $u \neq \vec{0}$ (property (5) of dots). So a = 0.
Similarly by dotting with v and w you can get b = 0 and c = 0.
So the only solution is a = b = c = 0. So u,v,w are ind.

method 2 Could one of u,v,w be a combination of the others. Suppose u = av + bw. Then

$u\cdot u = u(av + bw)$

$u\cdot u = a(u\cdot v) + v(u\cdot w) = a(0) + b(0) = 0$

But this is a contradiction because $u\cdot u$ can't be 0 since $u \neq \vec{0}$.
So you can't have u = av + bw .
Similarly v can't be a combination of u and w; and w can't be a comb of u and v.
So u,v,w are ind.

8. Not possible in R^3, R^4, R^9. Nonzero orthog vectors are ind and you can't have 10 ind vectors in R^3, R^4, R^9.

Possible in R^{10}; for instance here are 10 nonzero orthog vectors in R^{10}.

$$\begin{array}{l}(1,0,0,0,0,0,0,0,0,0)\\(0,1,0,0,0,0,0,0,0,0)\\ \vdots \\(0,0,0,0,0,0,0,0,1,0)\\(0,0,0,0,0,0,0,0,0,1)\end{array}$$

SOLUTIONS Section 2.3

1.(a) (i) It is supposed to be clear by inspection that u,v,w are independent. (For instance v can't be a combination of u and w because v has a nonzero second coord but v and w have zero second coords.) So u,v,w is a basis since any 3 ind vectors in R^3 are a basis.

(ii) By inspection, $x = -2u - \frac{3}{2}v - 4w$ so x has coords $-2, -\frac{3}{2}, -4$ w.r.t. basis u,v,w.

(b) (i) I'll show that p,q,r are ind. Then since there are 3 of them, they are a basis for R^3.

method 1 for showing they are ind (Use (4) from the preceding section.)
Look at the vectors q,r,p in that order (because it's convenient).

The vector q is not $\vec{0}$.
r is not a multiple of q.
And p can't be a combination of r and q because p has a nonzero first component but r and q have zero first components.
So q,r,p are ind.

method 2 for showing they are ind (Use (3) from the preceding section.)

Try to solve $ap + bq + cr = \vec{0}$. This is the system of equations

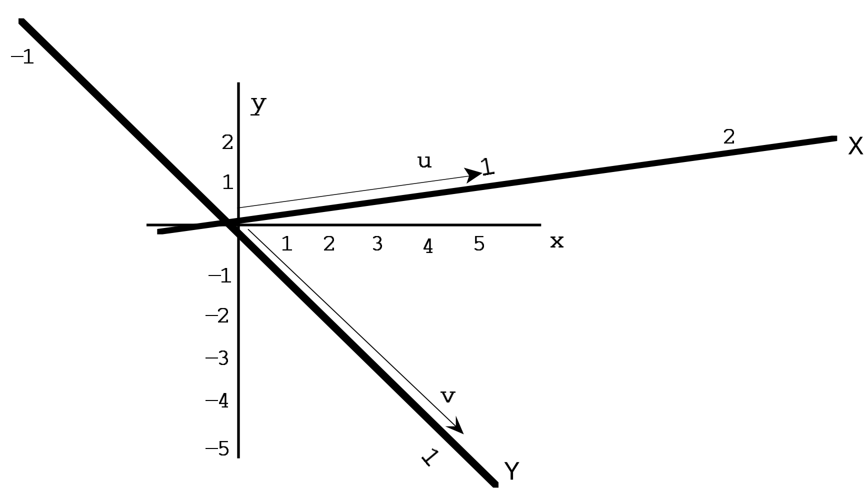
$$\begin{aligned} 1a + 0b + 0c &= 0 \\ 0a + 1b + 0c &= 0 \\ 0a + 1b + 1c &= 0 \end{aligned}$$

The only solution is a=0, b=0, c=0. So p,q,r must be ind.
(ii) Solve ap + bq + cr = x; i.e., solve

$$\begin{aligned} 1a + 0b + 0c &= 2 \\ 0a + 1b + 0c &= 3 \\ 0a + 1b + 1c &= 4 \end{aligned}$$

Solution is a = 2, b = 3, c = 1.
So x = 2p + 3q + r; x has coords 2,3,1 w.r.t. basis p,q,r.

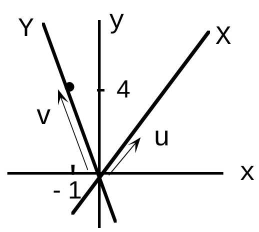
2. (a) The vectors u and v are independent since neither is a multiple of the other. And any 2 ind vectors in R^2 are a basis for R^2.
(b) See the diagram.



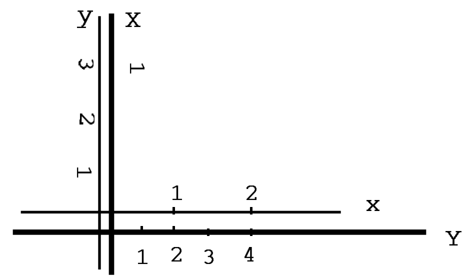
Problem 2(b)

3. u = 12i, v = 12j

4. The line y = 3x has slope 3. The vector (1,3) points along the line.
The first basis vector is $u = (1,3)_{\text{unit}} = (1/\sqrt{10},\ 3/\sqrt{10})$.
The second basis vector is v = (−1,4).



Problem 4



Problem 5

5. (a) The X-axis lies along the old y-axis (see the diagram).
The Y-axis lies along the old x-axis.
The scale on the X-axis is the yard.
The scale on the Y-axis is the half-foot.
(b) $X = \frac{1}{3}y$, $Y = 2x$.

6. (a) If a bunch of vectors is a basis for R^3 then every vector in R^3 can be written in terms of those vectors *in exactly one way*. In other words, they span "uniquely".

If all you know is that a bunch of vectors spans R^3, then you know that every vector in R^3 can be written in terms of vectors in the bunch but you do not have any guarantee that vectors in R^3 can be written in only one way in terms of them. (If you got that guarantee then the bunch of vectors would be promoted to being a basis.)

(b) One way to do it is to pick any basis and then toss in some extra vector(s). The new bunch still spans R^3 but isn't independent (because the extra vectors are combinations of the basis vectors). For instance, the vectors

i = (1,0,0)
j = (0,1,0)
k = (0,0,1)
p = (2,0,0)

span R^3 because every vector can be written in terms of them (in many ways actually) but they are not a basis (for one thing, they are not ind and for another, a basis for R^3 contains exactly 3 vectors, not 4).

Any bunch of more than 3 vectors which contains an independent threesome will span R^3 but not be a basis.

(c) No. A basis for R^3 must span R^3.

7.(a) I want to find a,b,c,d so that

au + bv + cw + dp
(3,0,5,6) = (a,0,0,0) + (b,0,b,0) + (0,0,0,c) + (2d,0,d,0)

$a + b + 2d = 3$
$b + d = 5$
$c = 6$

There are lots of solutions. One solution is c=6, d=0,b=5,a=−2.
Another solution is c=6, d=5,b=0, a=−7.
Another solution is c=6, d=1, b=4, a = −3.

So x = −2u + 5v + 6w.
And x = −7u + 6w + 5p
And x = −3u + 4b + 6w + p etc. (There are millions of answers.)
(b) The second component of y is not zero. But u,v,w,p all have zero second coords. That makes it impossible to write y as a combination of u,v,w,p.
(c) *answer 1* The vector x from part (a) can be written in more than one way in terms of u,v,w,p. So u,v,w,p can't be a basis.
answer 2 The vector y from part (b) can't be written at all in terms of u,v,w,p. So u,v,w,p can't be a basis.
answer 3 I noticed that p = u + v. So u,v,w,p are dep. So they can't be a basis.

8. (a) The solution says that you can pick any value for c. This means there is more than one way to write x as a combination of u,v,w,p. So u,v,w,p can't be a basis (there are 4 of them --- the right number --- but they are dep).
(b) x can't be written in terms of u,v,w,p, i.e., u,v,w,p don't span R^4. So u,v,w,p are not a basis (they must be dep).

9. (a) Can't find $p \cdot q$ or $\|q\|$.
(b) *method 1* You can use the u,v coords to dot and norm as you would use i,j coords.

$p \cdot q = 6 + 20 = 26, \quad \|q\| = \sqrt{9 + 25} = \sqrt{34}$

method 2 (the long way)

$$\begin{aligned} p \cdot q &= (2u + 4v)\cdot(3u + 5v) \\ &= 6u\cdot u + 20v\cdot v + 22u\cdot v \\ &= 6\|u\|^2 + 20\|v\|^2 + 22u\cdot v \\ &= 6(1) + 20(1) = 22(0) \text{ since u and v are orthonormal} \\ &= 2 + 20 = 26 \end{aligned}$$

$$\begin{aligned} \|q\|^2 = q\cdot q &= (3u + 5v)\cdot(3u + 5v) \\ &= 9u\cdot u + 25v\cdot v + 30u\cdot v \\ &= 9\|u\|^2 + 25\|v\|^2 + 30u\cdot v \\ &= 9(1)+ 25(1) + 30(0) \text{ since u and v are orthonormal} \\ &= 34 \end{aligned}$$

so $\|q\| = \sqrt{q\cdot q} = \sqrt{34}$

(c)
$$\begin{aligned} p \cdot q &= (2u + 4v)\cdot(3u + 5v) \\ &= 6u\cdot u + 20v\cdot v + 22u\cdot v \\ &= 6\|u\|^2 + 20\|v\|^2 + 22u\cdot v \\ &= 484 \end{aligned}$$

$$\begin{aligned} \|q\|^2 = q\cdot q &= (3u + 5v)\cdot(3u + 5v) \\ &= 9u\cdot u + 25v\cdot v + 30u\cdot v \\ &= 9\|u\|^2 + 25\|v\|^2 + 30u\cdot v \\ &= 631 \end{aligned}$$

so $\|q\| = \sqrt{631}$.

10. (a) $\mathbf{u}_{\text{unit}} = \dfrac{\mathbf{u}}{\|\mathbf{u}\|} = \dfrac{1}{2}\,\mathbf{u}$, $\mathbf{v}_{\text{unit}} = \dfrac{1}{3}\,\mathbf{v}$, $\mathbf{w}_{\text{unit}} = 4\mathbf{w}$

So $\mathbf{x} = 4\mathbf{u} + 5\mathbf{v} + 9\mathbf{w}$

$$= 4(2\mathbf{u}_{\text{unit}}) + 5(3\mathbf{v}_{\text{unit}}) + 9\left(\frac{1}{4}\,\mathbf{w}_{\text{unit}}\right)$$

$$= 8\,\mathbf{u}_{\text{unit}} + 15\,\mathbf{v}_{\text{unit}} + \frac{9}{4}\,\mathbf{w}_{\text{unit}}.$$

(b) The coords of $\mathbf{x}$ w.r.t. the orthonormal basis $\mathbf{u}_{\text{unit}}$, $\mathbf{v}_{\text{unit}}$, $\mathbf{w}_{\text{unit}}$ are 8, 15, 9/4.

So by (6), $\|\mathbf{x}\| = \sqrt{64 + 225 + \frac{81}{16}}$.

11. Yes. Here's why.
(1a) holds for these vectors, by hypothesis.
And by the unique representation rule in Section 2.2, when you do write a vector in terms of them, it can be done in only one way since they are ind. So (1b) holds also. So the vectors are a basis. (And now I know there must be six of them.)

SOLUTIONS 2.4

1. (a) Let $P = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$. Then $P^{-1} = \begin{bmatrix} 3/4 & -1/2 \\ -1/4 & 1/2 \end{bmatrix}$ and $P^{-1}\begin{bmatrix} 4 \\ 8 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

So (4,8) = −u + 3v. The new coords of the point (4,8) are −1,3.
(b) −u + 3v = −(2,1) + 3(2,3) = (4,8).

(c) $P^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \text{col 1} \\ \text{of} \\ P^{-1} \end{bmatrix}$. So $i = \frac{3}{4}u - \frac{1}{4}v$. Similarly $j = -\frac{1}{2}u + \frac{1}{2}v$.

Check: $\frac{3}{4}u - \frac{1}{4}v = \frac{3}{2}i + \frac{3}{4}j - \frac{1}{2}i - \frac{3}{4}j = i$

(d) $P^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix}$ so $X = \frac{3}{4}x - \frac{1}{2}y$ and $Y = -\frac{1}{4}x + \frac{1}{2}y$.

$P\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$ so x = 2X + 2Y , y = X + 3Y

2. (a) Solve ai + bj + cw = x to get the new coords a,b,c; i.e., solve the system

$$\begin{aligned} a \quad + 2c &= 4 \\ b + 3c &= 1 \\ 4c &= -2 \end{aligned}$$

Start from the bottom up to see that the solution is c = −1/2, b = 5/2, a = 5. So $x = 5i + \frac{5}{2}j - \frac{1}{2}w$.

Check: $5i + \frac{5}{2}j - \frac{1}{2}w = (5,0,0) + (0, \frac{5}{2}, 0) - (1, \frac{3}{2}, 2) = (4,1,-2)$ OK

(b) $P = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & -3/4 \\ 0 & 0 & 1/4 \end{bmatrix}$, $P^{-1}\begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ 5/2 \\ -1/2 \end{bmatrix}$

So the new coords of $\vec{x}$ are X = 5, Y = 5/2 , Z = −1/2.

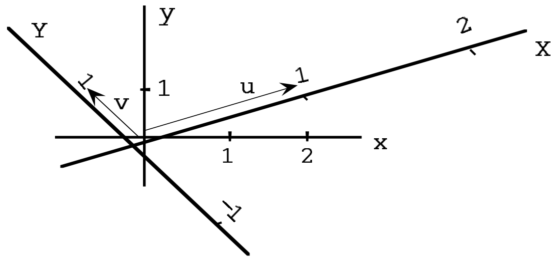
3. *answer 1* (most succinct) 2i + 3j + 4k = 5u + 6v + 7w.
answer 2 The vector with coordinates 2,3,4 w.r.t. basis i,j,k has coordinates 5,6,7 w.r.t. basis u,v,w.

4. The matrix B^{-1} converts from u,v,w coords to i,j,k coords. So B^{-1} is the standard basis changing matrix (the one I usually call P) and the columns of B^{-1} give the i,j,k coords of the vectors u,v,w. (You can stop here if you don't have a calculator that does inverses.)
I found that

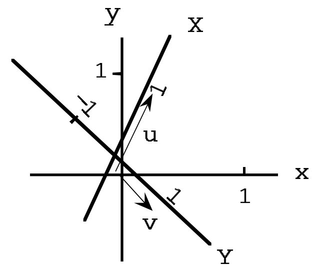
$$B^{-1} = \begin{bmatrix} -4 & -5 & 3 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{bmatrix}$$

So u = (−4,3,−1), v = (−5,3,−1), w = (3,−2,1).

5. (a) $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix}$ so $P = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$
The new basis is u = 2i + j, v = −i + j (the cols of P).



Problem 5a



Problem 5b

(b) $\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ so $P^{-1} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$

Invert to get $P = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{bmatrix}$. New basis is $u = (\frac{1}{3}, \frac{2}{3})$, $v = (\frac{1}{3}, -\frac{1}{3})$.

Question How come you had to invert a matrix in part (b) but not in part (a).
Answer In part (a), x and y were expressed in terms of X and Y. Turned out that this let you find P immediately. In part (b), X and Y were expressed in terms of x and y.

6.(a) The new coord system just involves a change of scale on the old axes.

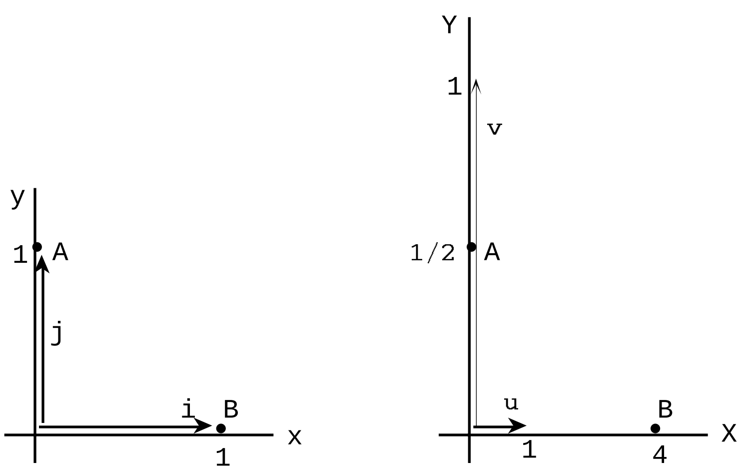
The scale on the X-axis is $\frac{1}{4}$ the scale on the x-axis.

The scale on the Y-axis is twice the scale on the y-axis.

The new basis vectors are $u = \frac{1}{4}\,i$, $v = 2j$.

(b) $\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ So $P^{-1} = \begin{bmatrix} 4 & 0 \\ 0 & 1/2 \end{bmatrix}$, $P = \begin{bmatrix} 1/4 & 0 \\ 0 & 2 \end{bmatrix}$

The new basis vectors are the columns of P, namely $u = \frac{1}{4}\,i$, $v = 2j$.



Old New

Problem 6(a)

7. The new coords of A are X = 2, Y = 6 (1 foot is 2 half-feet, 3 feet is 6 half-feet).

New basis is $u = \frac{1}{2}\,i$, $v = \frac{1}{2}\,j$.

8. (a) Only the scale changed. $u = \frac{1}{3}\,i$, $v = \frac{1}{2}\,j$

(b) Only the scale changed. $X = \frac{1}{4}\,x$, $Y = 3y$. So $x = 4X$, $y = \frac{1}{3}\,y$ and the circle has equation $(4X)^2 + (\frac{1}{3}y)^2 = 5$, $16X^2 + \frac{1}{9}Y^2 = 5$.

9. (a) Let $A = \begin{bmatrix} 4 & 2 \\ 5 & 3 \end{bmatrix}$. Then $A^{-1} = \begin{bmatrix} 3/2 & -1 \\ -5/2 & 2 \end{bmatrix}$ and A^{-1} converts from coords w.r.t. i,j to coords w.r.t. u,v. In other words, $A^{-1} \begin{bmatrix} \text{coords} \\ \text{w.r.t.} \\ \text{i and j} \end{bmatrix} = \begin{bmatrix} \text{coords} \\ \text{w.r.t.} \\ \text{u and v} \end{bmatrix}$

(b) $B = \begin{bmatrix} 6 & 8 \\ 7 & -1 \end{bmatrix}$ converts from coords w.r.t. p,q to coords w.r.t. i,j.

(c) Put (a) and (b) together. $A^{-1}B$ converts from p,q coords to u,v coords, i.e.,

$$A^{-1}\, B \begin{bmatrix} \text{coords} \\ \text{w.r.t} \\ \text{p and q} \end{bmatrix} = \begin{bmatrix} \text{coords} \\ \text{w.r.t.} \\ \text{u and v} \end{bmatrix}$$

(d) Answer here must be the inverse of the answer to part (c). So the answer is $(A^{-1}\, B)^{-1} = B^{-1}\, A$

10. It means that p = 2u + 3v.

11. (a) $P = \begin{bmatrix} 3 & 5 \\ 2 & 7 \end{bmatrix}$. Then $P \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$.
So the old coordinates of p are x=1, y=-3.
(b) p = 2u − v = (6,4) − (5,7) = (1,−3).
So the old coords of p are x=1, y= −3.

12. (a) $P = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & 1/2 \end{bmatrix}$, $P^{-1} \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$

So p = −2u + 3v

(b) The u,v basis is orthogonal so $p = \frac{u\cdot p}{u\cdot u}\, u + \frac{v\cdot p}{v\cdot v}\, v = \frac{-2}{1}\, u + \frac{12}{4}\, v = -2u + 3v$.

13.(a) First write p in terms of u and v; i.e., find the coords of p w.r.t. new basis u,v. You can do it with a system of equations but I'll use the basis changing matrix.

Let $P = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$. Then $P^{-1} = \frac{1}{-10} \begin{bmatrix} 2 & -4 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2/10 & 4/10 \\ 3/10 & -1/10 \end{bmatrix}$,

$P^{-1} \begin{bmatrix} 10 \\ 15 \end{bmatrix} = \begin{bmatrix} 4 \\ 3/2 \end{bmatrix}$ so $p = \underbrace{4u}_{q} + \underbrace{\tfrac{3}{2}v}_{r}$.

So $q = 4u = (4,12)$ and $r = \frac{3}{2}v = (6,3)$.

(b) t is the vector projection of p onto the vector v. Use the formula in Fig 7, Section 2.1:

$$t = \frac{p\cdot v}{v\cdot v}\, v = \frac{70}{20}\, v = \frac{7}{2}v = (14,7)$$

$$s = \frac{p\cdot u}{u\cdot u}\, u = \frac{55}{10}\, u = \left(\frac{55}{10}, \frac{165}{10}\right)$$

14. New basis is $u = (\cos 29^\circ, \sin 29^\circ)$, $v = (-\sin 29^\circ, \cos 29^\circ)$

method 1 (using the formula for finding coords w.r.t. an orthonormal basis)

$$p = (u\cdot p)u + (v\cdot p)v = (6 \cos 29^\circ + 7 \sin 29^\circ)u + (-6 \sin 29^\circ + 7 \cos 29^\circ)v$$

New coords are $X = 6 \cos 29^\circ + 7 \sin 29^\circ$, $Y = -6 \sin 29^\circ + 7 \cos 29^\circ$

method 2

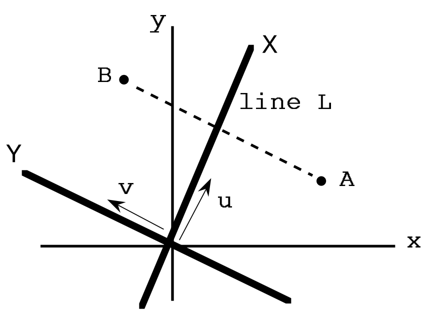
$P = \begin{bmatrix} \cos 29^\circ & -\sin 29^\circ \\ \sin 29^\circ & \cos 29^\circ \end{bmatrix}$, $|P| = 1$, $P^{-1} = \begin{bmatrix} \cos 29^\circ & \sin 29^\circ \\ -\sin 29^\circ & \cos 29^\circ \end{bmatrix}$,

$$P^{-1}\begin{bmatrix}6\\7\end{bmatrix} = \begin{bmatrix}6\cos 29^\circ + 7\sin 29^\circ\\ -6\sin 29^\circ + 7\cos 29^\circ\end{bmatrix}.$$

So the new coords are $X = 6\cos 29^\circ + 7\sin 29^\circ$, $Y = -6\sin 29^\circ + 7\cos 29^\circ$.

15. The vectors u,v,w,r are nonzero (because they don't have 0 norm, they have norm 1) and from §2.3, a set of 4 nonzero orthogonal vectors in R^4 must be a basis for R^4.
By (8), the numbers $u\cdot p$, $v\cdot p$, $w\cdot p$, $r\cdot p$ are the coords of p w.r.t. the orthonormal basis u,v,w,r.
So by (5) in Section 2.3, $\sqrt{(u\cdot p)^2 + (v\cdot p)^2 + (w\cdot p)^2 + (r\cdot p)^2}$ is $\|p\|$.

16. *step 1* Let $u = i + 3j$ (points along line L because the line and the vector both have slope 3; see the diagram) and let $v = -3i + j$ (perp to u).
It would also be OK to use $u = i + 3j$, $v = 3i - j$ or $u = -i - 3j$ and $v = -3i + j$ or $u = -i - 3j$, $v = 3i - j$.
It is also OK to normalize u and v and use u_{unit} and v_{unit} as your basis. In this problem it isn't necessary but it often is a good precaution to take.



Problem 16

step 2 $P = \begin{bmatrix}1 & -3\\ 3 & 1\end{bmatrix}$, $P^{-1} = \frac{1}{10}\begin{bmatrix}1 & 3\\ -3 & 1\end{bmatrix}$

$$P^{-1}\begin{bmatrix}a\\b\end{bmatrix} = \frac{1}{10}\begin{bmatrix}a+3b\\-3a+b\end{bmatrix}$$

So the new coords of A are $X = \frac{1}{10}a + \frac{3}{10}b$, $Y = -\frac{3}{10}a + \frac{1}{10}b$.

step 3 When you reflect a point in the X-axis in an orthogonal coord system all you do is change the sign of its Y-coord. That was the whole point of switching to a new coord system where the line L was an axis.
So the coords of point B, in the new system, are

$$X = \frac{1}{10}a + \frac{3}{10}b,\quad Y = \frac{3}{10}a - \frac{1}{10}b$$

step 4
$$P\begin{bmatrix}\frac{1}{10}a + \frac{3}{10}b\\ \frac{3}{10}a - \frac{1}{10}b\end{bmatrix} = \begin{bmatrix}-\frac{8}{10}a + \frac{6}{10}b\\ \frac{6}{10}a + \frac{8}{10}b\end{bmatrix}$$

The coords of B in the standard coord system are

$$x = -\frac{8}{10}a + \frac{6}{10}b,\quad y = \frac{6}{10}a + \frac{8}{10}b$$

And now you have a formula for reflecting an arbitrary point (a,b) in the line $y = 3x$. The reflection is the point $\left(-\frac{8}{10}a + \frac{6}{10}b,\ \frac{6}{10}a + \frac{8}{10}b\right)$.

17. Let $P = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 3 & 1 & 1 \end{bmatrix}$. Let $Q = \begin{bmatrix} \pi & 0 & 5 \\ 2 & 1 & 5 \\ \sqrt{3} & 2 & 5 \end{bmatrix}$.

P converts from u,v,w coords to i,j,k coords.
Q converts from p,q,r coords to i,j,k coords..
P^{-1} converts from i,j,k coords to u,v,w coords.
$P^{-1}Q$ converts from p,q,r coords to u,v,w coords (this is the answer).

18. (a) $u \cdot v = 0$, $u \cdot w = 0$, $v \cdot w = 0$ so u,v,w are orthogonal.
They are a basis because any set of 3 orthogonal vectors in R^3 is a basis for R^3.
(b) $x = \dfrac{u \cdot x}{u \cdot u}\, u + \dfrac{v \cdot x}{v \cdot v}\, v + \dfrac{w \cdot x}{w \cdot w}\, w = \dfrac{\pi+3}{2}\, u + \dfrac{-\pi+3}{2}\, v + 3w$

SOLUTIONS Section 2.5

1. (a)(i) The subspace consists of all multiples of u. It is 1-dim and u alone can be the basis.
(ii) p and q are not in the space since they are not multiples of u.
(iii) Any vector in the space can be the basis so a second basis consists of the vector 2u = (2,4,6,8,10), a third basis contains just πu = (π,2π,4π,6π,8π,10π) etc.
(b) (i) u and v are ind (because neither is a multiple of the other) so the subspace is 2-dim. A basis is u,v.
(ii) p = −2u + 2v so p is in the subspace.
Its coords w.r.t. the basis u,v are −2,2.
The vector equation au + bv = q boils down to the system of equations a+b=4, 2a+b=5, 3a=6, 4a=7, 5a=8. Thre are no solutions so q can't be written in the form au + bv. So q is not in the subspace.
(iii) Any two ind vectors in the subspace constitute a basis. So here are a few more bases.
basis 1 2u,v
basis 2 u, u + 2v = (3,4,3,4,5)
basis 3 u−v = (0,1,3,4,5), u+v = (2,3,3,4,5) etc.
(c) u,v,w are dep but u,v are ind (so are u,v and so are v,w). So u,v,w span the same space as u,v. Answers here are the same as in (b).

2. (a) The line through the origin in the direction of u, i.e., the line y = 3x.
(b) The x-axis.
(c) Every vector can be written in terms of u and i so the subspace is all of R^2 itself.
(d) All of R^2

3. (a) The x-axis.
(b) The plane through the origin determined by i and p.
(c) Same answer as (b). The vector i+p is a combination of i and p so the subspace spanned by i,p, and i+p is the same as the subspace spanned by i,p.
(d) The vectors are ind and in fact are a basis for R^3. So the subspace they span is R^3 itself.
(e) R^3

4. False. The best you can say is that the dimension is ≤ 5. It is 5 only if the vectors are independent.

5. (a) Dependent since the space they span is only 2-dim.
(b) Dep (every threesome is dependent) since the space they span is only 2−dim.
(c) Can't tell. At least one pair is ind since the four vectors span a 2-dim subspace but you can't tell which pair(s).
(d) Dependent. These three vectors are in the subspace since the subspace is spanned by u,v,w,p. But the subspace is 2-dim so any three vectors in the subspace must be dep.

6. A typical vector in the subspace is of the form $a\vec{u} + b\vec{v} + c\vec{w}$. I'll show that $\vec{x}$ is orthog to this typical vector. That will show that x is orthog to every vector in the subspace.

$$\begin{aligned} x\cdot(au + bv + cw) &= a(x\cdot u) + b(x\cdot v) + c(x\cdot w) && \text{dot rules}\\ &= a(0) + b(0) + c(0) && \text{x is orthog to u,v,w}\\ &= 0 \quad \text{QED} \end{aligned}$$

7. No. Once a subspace contains $\vec{u}$ it also contains $2\vec{u}$, $\pi\vec{u}$, $\sqrt{3}\,\vec{u}$ etc. since a subspace must be closed under scalar mult. So it must contain infinitely many vectors or contain just one vector if it is the trivial subspace containing only $\vec{0}$. A subspace contains either exactly one vector or infinitely many. But never just 7 vectors.
8. (a) No

reason 1 It isn't closed under addition. The vectors (0,0,0,0,2) and (1,2,3,4,2) are in the set but their sum is (1,2,3,4,4) which is not in the set.
reason 2 It isn't closed under scalar mult. The vector u = (0,0,0,0,1) is in the set but 3u = (0,0,0,0,3) is not in the set

reason 3 It doesn't contain $\vec{0}$ so it can't be a subspace
(b) No
reason 1 Not closed under addition. The vectors (1,1,1,1,1) and (2,4,8,16,32) are in the set but their sum is (3,5,9,17,33)) which is not in the set.
reason 2 Not closed under scalar mult. The vector u = (1,1,1,1,1) is in the set but 2u = (2,2,2,2,2) is not in the set
(c) Yes
reason 1 Closed under addition and scalar mult
Two typical points in the set are u = (a,b,2a,c,3b) and v = (d,e,2d,f,3e). Then

u+v = (a+d,e+b,2a+2d,c+f,3b+3e)

and

ku = (ka,kb,2ka,kc,3kb)

Both u+v and ku have the pattern $x_3 = 2x_1$ and $x_5 = 3x_2$ so they are in the set.

reason 2 The set consists of all combinations of the vectors
u = (1,0,1,0,0), v = (0,1,0,0,3), w = (0,0,0,1,0)
You can see this by inspection or you can get it like this:
The set consists of vectors $(x_1, x_2, x_3, x_4, x_5)$ where

$$\begin{aligned}
x_1 &= a &&= 1a + 0b + 0c\\
x_2 &= b &&= 0a + 1b + 0c\\
x_3 &= 2a &&= 2a + 0b + 0c\\
x_4 &= c &&= 0a + 0b + 1c\\
x_5 &= 3b &&= 0a + 3b + 0c
\end{aligned}$$

So

$$\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5 \end{bmatrix} = a\underbrace{\begin{bmatrix} 1\\0\\2\\0\\0 \end{bmatrix}}_{u} + b\underbrace{\begin{bmatrix} 0\\1\\0\\0\\3 \end{bmatrix}}_{v} + c\underbrace{\begin{bmatrix} 0\\0\\0\\1\\0 \end{bmatrix}}_{w}$$

This shows that the set is spanned by u,v,w. So it's a subspace.

To find a basis, continue with the work in reason 2. The spanning vectors
u = ($\boxed{1}$,$\boxed{0}$,2,$\boxed{0}$,0), v = ($\boxed{0}$,$\boxed{1}$,0,$\boxed{0}$,3), w = ($\boxed{0}$,$\boxed{0}$,0,$\boxed{1}$,0) are independent. No one vector can be a combination of the others because of their 1st, 2nd and 4th components. So u,v,w are a basis.

(d) No
reason 1 Not closed under addition. If u = (2,4,0,0,0) and v = (1,3,π,π,π) then u and v are in the set but u+v = (3,7, who cares) and it isn't in the set.
reason 2 Not closed under scalar mult. If u = (2,4,0,0,0) then u is in the set but 2u = (4,8,0,0,0) is not in the set.

reason 3 The set doesn't contain $\vec{0}$.

9. (a) If r=0,s,=0,t=0 then the corresponding point is (0,0,0,0,0).

If r=1,s=0,t=−1 then the corresponding point is (3, 6, −1, 1, 0). And so on.

(b) The parametric equations can be written in vector form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = r \underbrace{\begin{bmatrix} 2 \\ 8 \\ 0 \\ 1 \\ 1 \end{bmatrix}}_{u} + s \underbrace{\begin{bmatrix} 4 \\ 9 \\ 8 \\ -1 \\ 1 \end{bmatrix}}_{v} + t \underbrace{\begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{w}$$

So the set of points satisfying the parametric equations is the subspace of R^5 spanned by u,v,w.

(c) If u,v,w are ind then the subspace is 3-dim. If u,v,w are dep then the subspace is 2-dim because, by inspection, at least u,v are ind (neither is a multiple of the other). I tested u,v,w using Mathematica.

```
In[1]
u = {2,8,0,1,1}; v = {4,9,8,-1,1}; w = {-1,2,1,0,1};
Solve[a u + b v + c w ==0 ]
Out[1]
{{a -> 0, c -> 0, b -> 0}}
```

So u,v,w are ind and the subspace is 3-dim.

10. The space is spanned by u = (1,1,1,1,1) (i.e., the set consists of all multiples of u) so it's a subspace.

(You can also show it's a subspace by showing that it's closed under addition and scalar mult.)

It's a 1-dim subspace with basis u. (Any vector in the subspace can serve as a basis.)

The coordinate of v w.r.t. my basis is 3 since v = 3u.

11. (a) The vectors in R^2 are 2-tuples.

The vectors in a 2-dim subspace of R^9 are 9-tuples although they do have only 2 coords when you write them in terms of a basis for the subspace.

(b) The set of all combinations of u,v,p,q.

(c) *answer 1* Statement (2) says more than (1). It means that not only do u,v,w span V but they are independent. From (1) alone, you don't know if they are ind.

answer 2 Statement (2) means that not only can every vector in V be written in terms of u,v,w but it can be done in only one way.
When (1) holds, every vector in V can be written in terms of u,v,w but you don't know whether representations are unique or not.

(d) A subset is any old collection of vectors from R^5.

A subspace is a special collection of vectors. It's the span of some bunch of vectors. Equivalently, it's a set of vectors that is closed under addition and scalar mult.

(e) *answer 1* Not unless u,v,w,p are independent.

anwer 2 Not unless every vector in the subspace can not only be written in terms of u,v,w,p but can be written *uniquely* in terms of u,v,w,p.

(f) It means that q *can* be written as au + bv + cw + dp but there's only one way to do it.

12. (a) It would mean that if u and v are in the set then so are u−v and v−u.

(b) True. Suppose u and v are in the subspace. Remember that u−v = u + (−1)v.
We know that (−1)v is in the subspace since the subspace is closed under scalar mult.
And then u + (−1)v is in the subspace because the subspace is closed under addition.

So u − v is in the subspace.
Similarly, v−u is in the subspace. QED
(c) It would mean that if u is in the set then u_{unit} is also in the set.

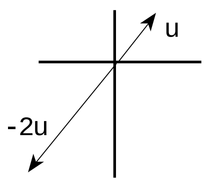
(d) True. $u_{unit} = \frac{u}{\|u\|} = \frac{1}{\|u\|} u$; So u_{unit} is a scalar multiple of u (the scalar is $1/\|u\|$). If u is in the subspace then so is u_{unit} because the subspace is closed under scalar mult.

(e) If u and v are vectors then u·v is a *scalar* which can't possibly be in the subspace (not unless this is all taking place in R^1). You can answer False or better still say that the question doesn't make much sense.

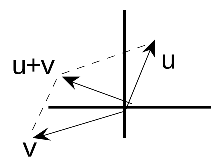
13. (a) u·v = 0, u·w = 0, v·w = 0

(b) $x = \frac{u \cdot x}{u \cdot u} u + \frac{v \cdot x}{v \cdot v} v + \frac{w \cdot x}{w \cdot w} w = \frac{10}{2} u + \frac{6}{1} v + \frac{7}{1} w = 5u + 6v + 7w$

14. (a) One possibility is the set of points in quadrant I
It's closed under addition and it's closed under mult by a *positive* scalar but if u is the vector in the diagram then u is in quad I but −2u isn't.
(b) One possibility is quadrants I and III put together. It's closed under scalar multiplication. But if u and v are the vectors in the diagram then u and v are *in* but u+v is *out*.



Problem 14(a)



Problem 14(b)

(c) One possibility is the set of points with all positive coords .
(d) One possibility is the set of points whose coords are either all positive or all negative.

15. Not a subspace.

reason 1 The subspace contains $\vec{0}$ so the leftovers don't contain $\vec{0}$.
So the leftovers can't be a subspace. This is true no matter what subspace you start with. The leftovers are never a subspace.
reason 2 The leftovers are not closed under addition. For instance (2,8,9,5) and (−1,π,2,−6) are leftovers but the sum (−1,8+π,7,−1) is not a leftover. (So even if you took the leftovers and added $\vec{0}$ to them to even have a chance, you still wouldn't get a subspace.)

16. The subspace is 3-dim. A cheap way to get a new basis is to essentially keep the 3 axes we have but change the scale, i.e., take various scalar multiples of the basis we have.

basis 1 (2,0,0,−1), (0,1,0,0), (0,0,1,0)

basis 2 $(\pi,0,0,-\pi)$, (0,3,0,0), $(0,0,\frac{1}{2},0)$

More generally, any 3 ind vectors in the space are a subspace

basis 3 (2,3,4,−2), $(\sqrt{3}, \frac{1}{3}, 6, -\sqrt{3})$, (0,2,0,0)

Here's a check that these vectors are ind since it isn't obvious

```
u = {2,3,4,-2}; v= {Sqrt[3],1/3,6,-Sqrt[3]};
w= {0,2,0,0};
Solve[a u + b v + c w == {0,0,0,0}]
{{c -> 0, b -> 0, a -> 0}}
```

etc.

17. The subspace is 3-dim (since u,v,w is a basis).
I'll show that the vectors u, v+w, v−w are ind.
And (very important), the vectors u, v+w, v−w are *in* the subspace since the subspace consists of all combinations of u,v,w.
That makes u, v+w, v−w a basis since any 3 ind vectors in a 3-dim subspace are a basis for the subspace.
Here's how to show that they are independent.

Look at the vector equation $a\vec{u} + b(\vec{v}+\vec{w}) + c(\vec{v}-\vec{w}) = \vec{0}$.

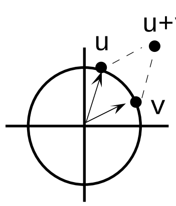
Rewrite it as $au + (b+c)v + (b-c)w = \vec{0}$.
The vectors u,v,w are ind so that means
(*) $a = 0,\ b+c = 0,\ b-c = 0$.
But the sol to the little system in (*) is $a = 0,\ b = 0,\ c = 0$.

So the only solution to $a\vec{u} + b(\vec{v}+\vec{w}) + c(\vec{v}-\vec{w}) = \vec{0}$ is $a=b=c=0$.
That makes u,v+w, v−w ind.
QED

18. A spanning set for the subspace.

19. (a) No.
reason 1 It's a rule that the only subspaces of R^2 are lines through the origin and R^2 itself. (If you start with some vectors and take all combinations you never end up with a circle; you always end up with a line through the origin or all of R^2.)
reason 2 The set of points on a circle is not closed under addition (or scalar mult). The diagram shows two points u and v on the circle but their sum is not on the circle.

reason 3 It doesn't contain $\vec{0}$.



Problem 19

20. I'll show that the set is closed under addition and scalar multiplication.
Suppose x and y are in the set. That means that
(*) $Ax = Bx$ and $Ay = By$

I want to show that x + y is in the set and kx is in the set. That means I want to show that
$A(x+y) = B(x+y)$ and $A(kx) = B(kx)$

Here's the argument.
$A(x+y) = Ax + Ay$ by matrix algebra
$= Bx + By$ by (*)
$= B(x+y)$ matrix algebra
and
$A(kx) = k(Ax)$ matrix algebra
$= k(Bx)$ by (*)
$= B(kx)$ matrix algebra again. QED

21. (a) Four equations in three unknowns:
$a + 5b + 6c = 4$
$2a + 6b + 8c = 4$

$$3a - 6b + 2c = -4$$
$$4a + 2b + 6c = -2$$

(b) The solution is $b = 1-c$, $a = -1-c$, c = anything.
There are infinitely many sols (pick any c). For example, one solution is $c = 1$, $b = 0$, $a = -2$. Another solution is $c = \pi$, $b = 1-\pi$, $a = -1-\pi$. In other words, $x = -2u + w$ and also $x = (-1-\pi)u + (1-\pi)v + \pi w$.
The fact that there is a solution at all means that x is in the subspace spanned by u,v,w.
The fact that there are many solutions means that x can be written in many ways in terms of u,v, w so u,v,w can't be a basis for the subspace they span; u,v,w must be dependent. The subspace is at most 2-dim. (I can see that u and v are ind since u is not a multiple of v, so the subspace *is* 2-dim.)

22. They all mean different things.
Statement (2) means that V contains all combinations of u,v,w and nothing else.; i.e., V consists [entirely] of all combinations of u,v,w; i.e., V is the set of all combinations of u,v,w [nothing more, nothing less].
Statement (3) means that V contains all combinations of u,v,w and could contain something else as well.
Statement (1) means that V contains some but not necessarily all combinations of u,v,w and nothing else; e.g., V could contain only the vector 2u+7w.

23. The set of vectors of the form $(x_1,x_2,x_3,x_4,0,0,0)$ is a subspace of R^7 (for one thing, it is closed under addition and scalar mult). In particular it is the 4-dim subspace spanned by $e_1 = (1,0,0,0,0,0,0)$, $e_2 = (0,1,0,0,0,0,0)$, $e_3 = (0,0,1,0,0,0,0)$, $e_4 = (0,0,0,1,0,0,0)$.
The vectors $u_1,\ldots, u_5$ are in the subspace. But 5 vectors in a 4-dim subspace must be dependent. So $u_1,\ldots, u_5$ are dependent.

SOLUTIONS review problems for Chapter 2

1. $w \cdot v$ is just a scalar so temporarily call it k for convenience. Then
$(v - kw) \cdot w = v \cdot w - (kw) \cdot w = v \cdot w - k(w \cdot w) = v \cdot w - k\|w\|^2 = v \cdot w - 9k$.
But $k = w \cdot v = v \cdot w$ so the answer is $-8(v \cdot w)$.

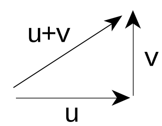
2. (a) False. As a counterexample let

$u = (1,0,0,0)$
$v = (0,1,0,0)$
$w = (2,3,0,5)$

They are ind but (1,0,0), (0,1,0), (2,3,0) are not ind.
(b) True.

3. (a) If $u \cdot v = 0$ then $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.
(See the diagram.)
(b) Given $u \cdot v = 0$. Then

$$\begin{aligned}\|u+v\|^2 &= (u+v)\cdot(u+v) = u\cdot u + 2u\cdot v + v\cdot v \\ &= u\cdot u + v\cdot v \\ &= \|u\|^2 + \|v\|^2 \quad \text{QED}\end{aligned}$$



Problem 3

4. Look at the subspace of R^{29} spanned by u,v,w. The dim of the subspace is 3 or less (exactly 3 if u,v,w are ind, less than 3 if they are dep).

The four vectors $2u + 6v$, $\pi v + \sqrt{5}\,w$, $7v + \frac{2}{3}w$, $8u - \pi v$ are in the subspace because they are all combinations of the generators u,v,w.

But 4 vectors in a 3-or-less-dimensional space must be dep (Section 2.5).

5. (a) Solve $au + bv + cw + dp = \vec{0}$ (this is a system of 29 equations in the 4 unknowns a,b,c,d).
If the only solution is $a = b = c = d = 0$ then the vectors are ind.
If there are other solutions in addition to the trivial one, then the vectors are dependent.
(b) Solve $au + bv + cw + dp = x$ (4 equations in the 4 unknowns a,b,c,d).
There has to be exactly one solution because a point has exactly one set of coordinates w.r.t. a basis. The values you get for a,b,c,d are the new coords of x.
(c) Solve $au + bv + cw = y$ (this is a system of 4 equations in 3 unknowns a,b,c).
If there are no solutions then y is *not* in the subspace. If there is at least one solution then y *is* in the subspace.
(d) Solve each of these six vector equations (each one is a system of 4 equations in 3 unknowns a,b,c):

$q = au + bv + cw$
$r = au + bv + cw$
$s = au + bv + cw$
$u = aq + br + cs$
$v = aq + br + cs$
$w = aq + br + cs$

If any one of the six systems has no sol then the two subspaces must be different.
Here's why. Suppose the system $r = au + bv + cw$ has no solution. Then r is not in the u,v,w subspace. But of course r is in the p,q,r subspace. So the two subspaces can't be the same
If all of the six systems have solutions then the two subspaces are the same.
Here's why. Since the first three systems have solutions, you know that q,r,s are combinations of u,v,w so they are in the u,v,w subspace. The subspace is closed under addition and scalar mult so every comb of q,r,s is also in the u,v,w subspace. So everything in the q,r,s subspace is also in the u,v,w subspace. Similarly, everything in the u,v,w subspace is also in the p,q,r subspace. So the two subspaces are the same.

(e) Let the purported inverse be

$$B = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix}$$

Solve the matrix equation AB = I. This amounts to 16 equations in 16 unknowns:

$$a_1 + 4b_1 + 3c_1 - d1 = 1$$
$$a_2 + 4b_2 + 3c_2 - d_2 = 0$$

etc.

If the system of equations has no solution then there is no inverse.
If it has a solution for the a's,b's,c's,d's then there is an inverse, and if you plug those values into B and you've got your inverse.

6. (iii) is right. Here's how to find the correct u and v.

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 5 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad P = \begin{bmatrix} 2 & 3 \\ 5 & -6 \end{bmatrix}^{-1} = \begin{bmatrix} 4/27 & 3/27 \\ 5/27 & -2/27 \end{bmatrix}$$

Basis is u = (4/27, 5/27), v = (3/27, −2/27) (the cols of P).

7. True *defense 1* Could 3v be a comb of 2u, 4w, 5p?

If 3v = a(2u) + b(4w) + c(5p) then $v = \frac{2}{3}a\,u + \frac{4}{3}b\,w + \frac{5}{3}c\,p$.

But that contradicts the hypothesis that u,v,w,p are ind.

Similarly, 2u can't be a comb of 3v, 4w, 5p and 4w can't be a comb of the others and 5p can't be a comb of the others.

defense 2 (neater) Look at the equation $a(2u) + b(3v) + c(4w) + d(5p) = \vec{0}$.
I want to show that the only way it can happen is with a = b = c = d = 0.

Rewrite the equation as $2a\,u + 3b\,v + 4c\,w + 5d\,p = \vec{0}$.
We know that u,v,w,p are ind. So we must have

$$2a = 0,\ 3b = 0,\ 4c = 0,\ 5d = 0$$

So

$$a = 0,\ b = 0,\ c = 0,\ d = 0$$

That makes u,v,w,p ind.

8. No. Here's an explanation by contradiction.
If 3u were in the subspace then all multiples of it would be in the subspace, including the multiple $\frac{1}{3}3u$. But $\frac{1}{3}3u$ is u which is *not* in the subspace, contradiction. So 3u can't be in it.

9. B is the basis changing matrix that converts from u,v,w coords to i,j,k coords. So the vector 3u − v + 2w is the same as the vector 9i + 13j + 13k. In other words, the coordinates of the vector (9,13,13) w.r.t. the basis u,v,w are 3,−1,2.

10. It's this system of 4 equations in 2 unknowns:

$$1 = 3a + 9b$$
$$2 = 4a + 10b$$
$$3 = 5a + 11b$$
$$4 = 6a + 12b$$

11. $P = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, $P^{-1} = \frac{1}{5}\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$

$P^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix}$ so $X = \frac{3}{5}x - \frac{1}{5}y,\ Y = -\frac{1}{5}x + \frac{2}{5}y$.

12. The set consists of all points with norm ≤ 1. It isn't a subspace.

reason 1 Not closed under addition.
If u = (1,0,0,0) and v = (0,1,0,0) then u and v are in the set (each has norm ≤ 1) but u + v is (1,1,0,0) which is not in the set (doesn't have norm ≤ 1; in particular, it has norm $\sqrt{2}$).

reason 2 Not closed under scalar multiplication.
If u = (1,0,0,0) then u is in the set (has norm ≤ 1) but 2u is not in the set (doesn't have norm ≤ 1; in particular, it has norm 2).

SOLUTIONS Section 3.1

1. (a) In the echelon form, col 3 = 2 col 1 + 3 col 2 (also col 5 = 2 col 4). Same is true of the original cols. So the five cols of M are a dep set. In the original M, col 3 = 2 col 1 + 3 col 2.
(b) Maximal number of ind rows in M is 3 since there are 3 nonzero echelon rows. So the set of 5 rows in M is dep. Can't find a dependency relation among the rows from the given info.
(c) The nonzero echelon rows p = (1,0,2,0,0), q = (0,1,3,0,0), r = (0,0,0,1,2) are a basis for the echelon row space and for the original row space.
(d) Original cols 1,2,4 are a basis (not the echelon cols).
(e) (i) No. x isn't a combination of p,q,r from part (c). The only combination that has a chance is 2p + 3q + 5r. This is the only combination that has the right 1st, 2nd and 4th coordinates but it doesn't have the right 3rd component. So x is not in the row space of M.
(ii) Can't tell whether x is in the column space of M.
warning Here's a *wrong* way to do it.
Tack column x onto the echelon form as a 6th col like this

$$\begin{matrix} 1 & 0 & 2 & 0 & 0 & 2 \\ 0 & 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 6 \end{matrix}$$

The 6th col is not a combination of the cols with pivots. So x is not a combination of the cols of M. So x is not in the col space,.

The correct version is to *row operate* the column vector x before tacking it on to the echelon form, using the same row operations that were used on M to get echelon form in the first place. Then you could draw a conclusion about the *row-operated-on x* and the echelon cols of M which carries back to the x and the original cols of M.
But you can't do that here because you don't know what row operation were used so you're stuck.
(f) (i) Yes. The nonzero echelon rows p = (1,0,2,0,0), q = (0,1,3,0,0), r = (0,0,0,1,2) are a basis for the original row space. y is in the row space because y = 5p − 3q. Since y in the row space it must be a combination of the original rows of M. But I don't know *what* combination it is----it is *not* necessarily 5 row 1 − 3 row 2.
(ii) Can' t tell with the given info. (same warning as in part (e)(ii) applies).
(g) Can't tell. The row space is 3-dimensional so there exists a set of 3 ind rows but you can't tell where it is. The fact that the first 3 echelon rows are ind does not carry back to the original rows.
(h) Dep. The row space of M is 3-dim so any 4 vectors in the row space are dependent.
(i) Dep. The corresponding echelon cols are dep (col 3 = col 1 + 3 col 2) and this carries back to the original cols.
(j) Ind. The corresponding echelon cols are ind and this carries back to the original cols.

2. (a) The col space is a 2-dim subspace (because there are 2 echelon cols with pivots) of R^4 (because the cols are 4-tuples). A basis is cols 1 and 4 of the original matrix.
The row space is a 2-dim space (because there are 2 nonzero echelon rows) of R^4 (because the rows are 4-tuples). The two nonzero echelon rows (1,2,3,0), (0,0,0,1) are a basis for the original row space.
(b) The col space is a 3-dim subspace of R^3. That means the col space *is* R^3 itself. So a basis is i,j,k. Another basis is cols 1,3,5 of the original matrix.
The row space is a 3-dim subspace of R^5. Since all the echelon rows are nonzero, they are a basis for the original row space. Another basis is the set of original rows.

(c) The col space and the row space are each 3-dim subspaces of R^3. So each *is* R^3. A basis of course is i,j,k. Another basis is the 3 original cols. Another is the 3 original rows.

3. Do these row ops.

row 2 = −2 row 1 + row 2
row 3 = −2 row 1 + row 3
row 3 ↔ row 2
row 2 = − row 2
row 1 = −row 2 + row 1
row 3 = $-\frac{1}{6}$ row 3
row 1 = −2 row 3 + row 1

$$\text{echelon form of } M = \begin{matrix} 1 & 2 & 3 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1/6 \end{matrix}$$

(a) The col space is a 3-dim subspace of R^3 with basis p = (1,2,2), q = (1,2,1), r = (2,−2,4), the original cols corresponding to the echelon cols with pivots.

Since the only 3-dim subspace of R^3 is R^3 itself, the col space is R^3 and the most sensible basis for the col space is i,j,k

(b) The row space is a 3-dim subspace of R^6 with basis u = (1,2,3,0,0, 2/3) v = (0,0,0,1,0,−1), w = (0,0,0,0,1, 1/6)

(c) Of course. The col space is all of R^3 so every 3-tuple is in it.

(d) The row space has basis

u = (1,2,3,0,0, 2/3), v =(0,0,0,1,0,−1), w = (0,0,0,0,1, 1/6).

By inspection, y = v + 12w. So y is in the row space.

4. Line up the vectors as cols and row op.

divide row 1 by 2
add −2 row 1 to row 2
add 4 row 1 to row 3
add −5 row 1 to row 4
add 4 row 2 to row 3
add −4 row 2 to row 4

Echelon form is

$$\begin{matrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$$

(a) In the echelon form, col 3 = 3 col 1 − 2 col 2. So u,v,w are dep and in particular, w = 3u − 2v

(b) The space spanned by u,v,w is 2-dim since there are two echelon cols with pivots. A basis is u,v, the original cols corresponding to the ech cols with pivots.

Another basis is u,w since in the echelon form, cols 1 and 3 are ind.

Another basis is v,w since in the echelon form, cols 2 and 3 are ind.

(c) I'm using method 1 because then I can take advantage of the work done in part (a). Take the row ops that turned the matrix with cols u,v,w into echelon form in part (a) and do them to x. In other words, row op on the matrix with cols u,v,w,x (could leave out w since we already know that u,v alone are a basis):

$$\begin{matrix} 1 & 0 & 3 & 3/2 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 15 \\ 0 & 0 & 0 & -11/2 \end{matrix}$$

This isn't echelon form but it's enough to see that in this new matrix, col 4 is not a combination of the first three cols. So back in the original, x is not a combination of u,v,w so x is not in the subspace spanned by u,v,w.

5.(a) *method 1* In A, col 3 = 2 col 1 + 3 col 2. So u_1, u_2, u_3, u_4 are dep. In particular $u_3 = 2u_1 + 3u_2$.

method 2 In B, there are maximally 3 ind rows. So u_1, u_2, u_3, u_4 are dep. But with this method I can't get a dependence relation.

(b) Cols 1,2,4 in A have pivots. So u_1, u_2, u_4 are a basis for the subspace spanned by u_1, u_2, u_3, u_4. Cols 1,3,4 in A and cols 2,3,4 in A are also ind so u_1,u_3,u_4 is another basis and so is u_2,u_3,u_4.

(c) Take the nonzero rows of B. The standard basis is

$v_1 = (1,0,0,4)$, $v_2 = (0,1,0,0)$, $v_3 = (0,0,1,1)$

(d) In A, col 4 is not a comb of the preceding cols. So u_4 is not a comb of u_1,u_2,u_3. So u_4 is not in the subspace spanned by u_1,u_2,u_3.

(e) In A, col 3 is a comb of the preceding cols. So u_3 is a comb of u_1,u_2. In particular, $u_3 = 2u_1 + 3u_2$. So u_3 is in the subspace spanned by u_1,u_2.

And u_1,u_2 are a basis for the subspace they span since they are independent.

The coords of u_3 w.r.t. basis u_1, u_2 are 2,3.

(f) From part (e), $u_3 = 2u_1 + 3u_2$. So $u_1 = \frac{3}{2}u_2 + \frac{1}{2}u_3$. So u_1 is in the subspace spanned by u_2 and u_3.

(g) The subspace spanned by u_1, u_2, u_3, u_4 has standard basis

$v_1 = (1,0,0,4)$, $v_2 = (0,1,0,0)$, $v_3 = (0,0,1,1)$.

By inspection, you need $2v_1 + 3v_2$ to get the first two components of p. But $2v_1+3v_2$ doesn't have the right last two components. So p is not in the subspace.

6. If you do any row ops here you are doing it the hard way (and people will laugh at you).

(a) The three rows span the 1-dimensional subspace of R^4 with basis (1,1,1,1). The subspace consists of all multiples of (1,1,1,1).

(b) The cols are vectors in R^2. The first two cols are independent since neither is a multiple of the other. So the cols span all of R^2; i.e., the col space *is* R^2. The nicest basis is i,j.

7. (a) Here's one method.Line up u,v,w as rows and row op into echelon form.
And line up p,q,r,s as rows and row op into echelon form.
The echelon forms respectively are

$$\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 1 \end{matrix} \qquad \text{and} \qquad \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix}$$

They have the same nonzero echelon rows. So the subspace spanned by u,v,w and the subspace spanned by p,q,r,s each has standard basis (1,0,0,0), (0,1, 3/2, 0), (0,0,0,1). So they are the same subspace.

(b) Try to solve the four systems of equations

$$
\begin{aligned}
p &= au + bv + cw \\
q &= au + bv + cw \\
(*) \quad r &= au + bv + cw \\
s &= au + bv + cw.
\end{aligned}
$$

Note that each of these vector equations is a system of 4 equations in 3 unknowns.
If say the system r = au + bv + cw has no solution then r is not in the subspace spanned by u,v,w so the subspace spanned by u,v,w can't be the same as the subspace spanned by p,q,r,s.
If each of the four systems has a solution (at least one) then you know that p,q,r,s are in the subspace spanned by u,v,w. That space is closed under addition and scalar mult so every combination of p,q,r,s is in the space spanned by u,v,w. So the subspace spanned by p,q,r,s is a subset of the space spanned by u,v,w.

Then solve the three systems of equations

$$
\begin{aligned}
u &= ap + bq + cr + ds \\
(**) \quad v &= ap + bq + cr + ds \\
w &= ap + bq + cr + ds
\end{aligned}
$$

Note that each of these vector equations is a system of 4 equations in 4 unknowns.
Again, if one of the systems has no solution then the two subspaces can't be the same.
If each of these three systems has at least one solution then the subspace spanned by u,v,w is a subset of the space spanned by p,q,r,s.

When two sets are each subsets of the other, they are the same set.
So here's the final rule: If any of the seven systems of equations displayed above has no solution then the two subspaces are not the same. If each of the seven systems has at last one solution then the two subspaces are the same. QED

8. (a) The five rows of M are ind. So the row space is a 5-dim subspace of R^5. But the only 5-dim subspace of R^5 is R^5 itself. So the row space *is* R^5. So every 5-tuple is in it, including u.
(b) The rows of M are dep. The row space is not all of R^5. Can't tell from this if u is in the row space or not.

9. Not possible to do it with a *square* matrix because for squares, having ind rows goes hand in hand with having ind cols.
But you can do it with a non–square matrix. Let

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 & 5 \\ 1 & 2 & 7 & 8 & 9 \end{bmatrix}$$

The rows are ind because neither is a multiple of the other. The cols are dep because 5 vectors in R^2 must be dep.

10. No. "Iff" means "if and only if".

11. According to the invertible list, a det is *non*zero iff its rows are ind.
So a det is *zero* iff its rows are dep. So the complete list of things to look for to spot a zero det should contain this one item:

some row is a combination of the other rows.

This one item is enough. The three things on the incomplete list are all special cases of this one item so you don't have to list them separately.

12. (a) Here's one method.
The subspace has basis u,v,w so it's 3−dim. See if p,q,r span the same subspace as u,v,w. If so, they are a basis for it.
Line up u,v,w as rows and row op into echelon form. The nonzero echelon rows are the standard basis for the subspace with basis u,v,w.
Then line up p,q,r as rows and row op into echelon form. The nonzero echelon rows are the standard basis for the subspace spanned by p,q,r.
If the two echelon forms are the same then p,q,r spans the same subspace as u,v,w and the student's answer is right. If the two echelon forms are not the same then the student's answer is wrong.
I got this echelon form both times:

$$\begin{matrix} 1 & 0 & 0 & 35/26 & -1/2 \\ 0 & 1 & 0 & -37/52 & 1/4 \\ 0 & 0 & 1 & 9/52 & 3/4 \end{matrix}$$

So her answer is right.
(b) I'll show that p,q,r are *in* the subspace and are ind. Then, since the subspace is 3-dim, that would make p,q,r a basis, in addition to u,v,w.

Line up u,v,w,p,q,r as cols and row op into echelon form. You should be able to use the echelon form to tell whether p,q,r are combinations of u,v,w. In particular, the echelon form turns out to be

$$\begin{matrix} 1 & 0 & 0 & 2 & 2 & -1 \\ 0 & 1 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

You can see that echelon cols 4,5,6 are combinations of echelon cols 1,2,3. So each of p,q,r is a comb of u,v,w so p,q,r are in the subspace.
Then start over again with just the columns p,q,r. This matrix row ops to I. So p,q,r are ind.
So p,q,r are 3 ind vectors in the 3-dim subspace. So they are a basis for the subspace.
Question How come you had to take that extra step in part (b) and show that p,q,r are ind but not in part (a).
Answer In the first part of (b), I showed that p,q,r are *in* the subspace. But three vectors in a 3-dim subspace are not necessarily a basis. They will be a basis only if they are ind. So I had to check. Otherwise the student could have gotten away with the answer (2,8,0,−3,1), (−2,−8,0,3,−1), (4,16,0,−6,2) which are *in* the subspace but not ind so not a basis.
On the other hand, in part (a), I did more than just show that p,q,r are *in* the 3-dim subspace; I showed that they actually *span* it. There's no way they could do that if they weren't independent. If you don't believe me and still want to check that p,q,r are ind, it doesn't take any extra work: they are ind because when p,q,r were lined up as rows, the echelon rows were all nonzero.

13. (a) Line up u,v,w as rows and row op to echelon form. Comes out to be

$$\begin{matrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{matrix}$$

The standard basis for the subspace spanned by u,v,w is
r = (1,2,0,3,0), s = (0,0,1,4,0), t = (0,0,0,0,1).
Points in the subspace are of the form

$$ar + bs + ct = (a,\ 2a,\ b,\ 3a+4b,\ c)$$

So $(x_1, x_2, x_3, x_4, x_5)$ is in the subspace iff it satisfies the equations

$x_2 = 2x_1$

$x_4 = 3x_1 + 4x_3$.

(b) The point $(2,4,-1,2,\pi)$ is in the subspace spanned by u,v,w because it satisfies the equations in (b). So it must be a combination of u,v,w.
With this method, I don't know *what* combination of u,v,w it is, but the problem didn't ask that.
(With this method, I do know how to write the point as a combination of r,s,t; it's $2r-s+\pi t$.)

SOLUTIONS Section 3.2

1.(a) The subspace is the x,y plane. The projection of (4,3,9) on the x,y plane is (4,3,0).

(b) Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 2 & 7 \\ 7 & 29 \end{bmatrix}$, $A^T\vec{q} = \begin{bmatrix} 7 \\ 23 \end{bmatrix}$.

The normal equations are

$$\begin{bmatrix} 2 & 7 \\ 7 & 29 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 7 \\ 23 \end{bmatrix}$$

$$\begin{aligned} 2a + 7b &= 7 \\ 7a + 29b &= 23 \end{aligned}$$

Solution is $a = 14/3$, $b = -1/3$. So the projection is $\frac{14}{3}u - \frac{1}{3}v = (4,3,0)$ again. The projection has coordinates $14/3$, $-1/3$ w.r.t. basis u,v for the subspace and it has coords 4,3,0 w.r.t. basis i,j,k for R^3.

warning The projection is $\frac{14}{3}u - \frac{1}{3}v$ and it is also (4,3,0). It is *not* $(\frac{14}{3}, -\frac{1}{3})$.

2. (a) Let $A = \begin{bmatrix} 2 & 0 \\ -1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 6 & -2 \\ -2 & 5 \end{bmatrix}$, $A^Tx = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$

The normal equations are

$$\begin{aligned} 6a - 2b &= 1 \\ -2a + 5b &= 5 \end{aligned}$$

Solution is $a = 15/26$, $b = 16/13$. So $x_{proj} = \frac{15}{26}u + \frac{16}{13}v$. In the original i,j,k coord system, $x_{proj} = \frac{15}{26}(2,-1,1,0) + \frac{16}{13}(0,2,0,1) = (\frac{15}{13}, \frac{49}{26}, \frac{15}{26}, \frac{16}{13})$.

(b) It isn't x_{proj} that is perp to something, it's $x - x_{proj}$ which is supposed to be perp to u and v (and hence to everything in the subspace spanned by u and v).

$$x - x_{proj} = (\frac{12}{17}, \frac{41}{34}, \frac{29}{34}, \frac{1}{17})$$

$$(x - x_{proj}) \cdot u = -\frac{4}{13} - \frac{3}{26} + \frac{11}{26} = 0$$

$$(x - x_{proj}) \cdot v = \frac{6}{26} - \frac{3}{13} = 0 \quad \text{Checks out.}$$

3. Find the projection of y onto the subspace spanned by u,v.

Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $A^T\vec{y} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

Normal equations are

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$\begin{aligned} 2a + b &= 0 \\ a + 2b &= 2 \end{aligned}$$

Solution is $a = -2/3$, $b = 4/3$. So $y_{proj} = -\frac{2}{3}u + \frac{4}{3}v$; this is the vector of the form $au + bv$ that is closest to y.

4. u,v are orthogonal so

$$q_{proj} = \frac{u \cdot q}{u \cdot u}u + \frac{v \cdot q}{v \cdot v}v = \frac{20}{29}u + \frac{-2}{20}v = \left(\frac{40}{29} - \frac{8}{20}, \frac{60}{29}, \frac{80}{29} + \frac{4}{20}\right) = \left(\frac{142}{145}, \frac{60}{29}, \frac{429}{145}\right)$$

The coords of the projection are $\frac{20}{29}, -\frac{2}{20}$ w.r.t. basis u,v for the subspace.

The coords of the projection are $\frac{142}{145}, \frac{60}{29}, \frac{429}{145}$ w.r.t. basis i,j,k for R^3.

5. $u_{proj} = \frac{v \cdot u}{v \cdot v}v = \frac{12}{31}v = \left(\frac{12}{31}, \frac{60}{31}, -\frac{24}{31}, \frac{12}{31}\right)$

6. (a) p and q are orthonormal so

$$u_p = (p \cdot u)p + (q \cdot u)q = p + 3q = \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right) + 3\left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right) = \left(-\frac{1}{3}, \frac{8}{3}, \frac{5}{3}\right)$$

The coords of the projection w.r.t. basis p,q are 1,3.

The coords of the projection w.r.t. basis i,j,k are $-\frac{1}{3}, \frac{8}{3}, \frac{5}{3}$.

(b) Let $A = \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \\ -1/3 & 2/3 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, the normal equations are

$$A^T A \begin{bmatrix} a \\ b \end{bmatrix} = A^T u = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

So solve

$$\begin{aligned} a + 0b &= 1 \\ 0a + b &= 3 \end{aligned}$$

Solution is a=1, b=3. So $u_{proj} = p + 3q$ as in part (a).

(c) A typical point in the subspace is $ap+bq = \left(\frac{2}{3}a - \frac{1}{3}b, \frac{2}{3}a + \frac{2}{3}b, -\frac{1}{3}a + \frac{2}{3}b\right)$.

The distance from this typical point to u is

$$\sqrt{\left(\frac{2}{3}a - \frac{1}{3}b - 1\right)^2 + \left(\frac{2}{3}a + \frac{2}{3}b - 2\right)^2 + \left(-\frac{1}{3}a + \frac{2}{3}b - 3\right)^2}$$

Find a and b that make this function minimum. It's sufficient to find a and b that make the radicand minimum (i.e., ignore the square root). Let

$$g = \left(\frac{2}{3}a - \frac{1}{3}b - 1\right)^2 + \left(\frac{2}{3}a + \frac{2}{3}b - 2\right)^2 + \left(-\frac{1}{3}a + \frac{2}{3}b - 3\right)^2$$

Find the partials and set them equal to 0:

$$\frac{dg}{da} = 2\left(\frac{2}{3}a - \frac{1}{3}b - 1\right) \cdot \frac{2}{3} + 2\left(\frac{2}{3}a + \frac{2}{3}b - 2\right) \cdot \frac{2}{3} + 2\left(-\frac{1}{3}a + \frac{2}{3}b - 3\right) \cdot -\frac{1}{3}$$

$$= 2a - 2$$

$$\frac{dg}{db} = 2b - 6$$

The solution to the system

$$\begin{aligned} 2a - 2 &= 0 \\ 2b - 6 &= 0 \end{aligned}$$

is $a = 1$, $b = 3$.
Answer is $p + 3q$ as in part (a).

7. (a) The line is a 1-dim subspace of R^2. The arrow $u = (1,3)$ points along the line so you can use u as the basis vector. Let $\vec{q} = (x,y)$. Then

$$\text{projection of } (x,y) = q_{proj} = \frac{u\cdot q}{u\cdot u}\,u = \frac{x+3y}{10}\,u = \left(\frac{1}{10}x + \frac{3}{10}y,\ \frac{3}{10}x + \frac{9}{10}y\right)$$

(b) The line $y = 3x + 7$ isn't a subspace (doesn't go through the origin). This section is about projecting into a subspace.

8. $(u\cdot x)u + (v\cdot x)v + (w\cdot x)w$ is the projection x_{proj} of x into the subspace with orthonormal basis u,v,w.

So $u\cdot x$, $v\cdot x$, $w\cdot x$ are the coords of x_{proj} w.r.t.an orthonormal basis.

By (6) in Section 2.3, When these components are squared, added, and the sum is square-rooted, you get the norm of x_{proj}.

So k is $\|x_{proj}\|$.

In R^3 it is geometrically clear that the length of a projection is $\le$ the length of the original vector (equal to it in the special case that the vector is its own projection). Predict that the same thing holds in R^{17}.So predict $k \le \|x\|$.

9. (a) The set of vectors of the form (a,2a,b,3b) is a subspace of R^4 and we want the projection of $q = (2,5,3,4)$ onto the subspace. (If the set of vectors wasn't a subspace, you can't do the problem using porjections.) To do this, you need a basis for the subspace. One basis is $u = (1,2,0,0)$, $v = (0,0,1,3)$.
It happens to be an orthog basis so

$$\text{closest point} = q_{proj} = \frac{u\cdot q}{u\cdot u}\,u + \frac{v\cdot q}{v\cdot v}\,v = \frac{12}{5}\,u + \frac{15}{10}\,v = \left(\frac{12}{5}, \frac{24}{5}, \frac{3}{2}, \frac{9}{2}\right)$$

(b) *answer 1* It means that the distance between q and $\left(\frac{12}{5}, \frac{24}{5}, \frac{3}{2}, \frac{9}{2}\right)$ is less than the distance between q and any other point of the form (a,2a,b,3b), i.e., the minimum value of $\sqrt{(2-a)^2 + (5-2a)^2 + (3-b)^2 + (4-3b)^2}$ occurs when $a = 12/5$ and $b = 3/2$.

answer 2 It means that the minimum value of $\|q - (a,2a,b,3b)\|$ occurs when (a,2a,b,3b) is the point $\left(\frac{12}{5}, \frac{24}{5}, \frac{3}{2}, \frac{9}{2}\right)$.

10. (a) u_1, u_2 u_3, u_4 is an orthogonal basis for R^4 (four orthog vectors in R^4 must be a basis). The expression is just x itself written in terms of the orthogonal basis.

(b) *answer 1* This is the projection of x into the subspace of R^4 with (orthogonal) basis $\vec{u}_2$, $\vec{u}_3$, $\vec{u}_4$. (It might actually equal x if x happened to be in the subspace.)

answer 2 It's the combination of u_2, u_3, u_4 that is closest to x; i.e., it's the best approximation to x of the form $a\vec{u}_2 + b\vec{u}_3 + c\vec{u}_4$.

11. (a) It's the projection of x into the subspace of R^5 with (orthogonal) basis $\overrightarrow{u_1}$, $\overrightarrow{u_2}$, $\overrightarrow{u_3}$, $\overrightarrow{u_4}$. (If x happens to be in that subspace then it is actually x itself.)

(b) The vector in part (a) is the best approximation to $\vec{x}$ of the form $au_1 + bu_2 + cu_3 + du_4$.

(c) Of all vectors in the subspace spanned by (with basis) u_1, u_2, u_3, u_4, the vector in part (a) is the one closest to x.

12. (a) V is a subspace of R^5 with basis

$u_1 = (0,1,0,0,0)$

$u_2 = (0,0,1,0,0)$

$u_3 = (1,0,0,1,0)$

$u_4 = (1,0,0,0,1)$

If you can't get this by inspection then here is a mechanical way: The set consists of all vectors of the form (x_1,x_2,x_3,x_4,x_5) where x_2,x_3,x_4,x_5 can be anything and $x_1 = x_4 + x_5$; in particular,

$$\begin{aligned} x_1 &= c+d \\ x_2 &= a \\ x_3 &= b \\ x_4 &= c \\ x_5 &= d \end{aligned}$$

Rewrite this as

$$\begin{aligned} x_1 &= 0a + 0b + 1c + 1d \\ x_2 &= 1a + 0b + 0c + 0d \\ x_3 &= 0a + 1b + 0c + 0d \\ x_4 &= 0a + 0b + 1c + 0d \\ x_5 &= 0a + 0b + 0c + 1d \end{aligned}$$

and then rewrite again:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = a \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{u_1} + b \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{u_2} + c \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{u_3} + d \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}}_{u_4}$$

This shows that the set of points consists of all combinations of u_1, u_2, u_3, u_4.The vectors are ind so they are a basis.

The vector in V closest to q is q_{proj}.

Let $A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Then $A^Tq = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 6 \end{bmatrix}$, $A^TA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

I used Mathematica to solve $A^TA \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 6 \end{bmatrix}$ and got a = 2, b = 3, c = 4/3, d = 7/3.

Answer is $q_{proj} = au_1 + bu_2 + cu_3 + du_4 = (\frac{11}{3}, 2, 3, \frac{4}{3}, \frac{7}{3})$.

(b) $q = (1,2,3,4,5)$ and $q_{proj} = (\frac{11}{3}, 2, 3, \frac{4}{3}, \frac{7}{3})$. The distance between them is

$$\sqrt{(\frac{8}{3})^2 + 0^2 + 0^2 + (\frac{8}{3})^2 + (\frac{8}{3})^2} = \frac{8}{3}\sqrt{3} \approx 4.6$$

Another vector in V is u_3. Distance from q to u_3 is $\sqrt{47} \approx 6.9$

Another vector in V is (5,2,2,0,5). Its distance from q is $\sqrt{33} \approx 5.7$
q is closer to q_{proj} than it is to these other vectors.

13. First of all, the plane goes through the origin (because the point (0,0,0) satisfies the equation). So it is a subspace of R^3 (very important because this section is about projecting onto *subspaces*).
Let $q = (0,2,-1)$. I want the projection of q onto the plane. Then the distance from point q to the plane (meaning the shortest distance) is $\|q-q_{proj}\|$.

To get the projection, I want a basis for the subspace.
Any two independent vectors in the plane will be a basis.
Let $u = (1,0,-2)$ and $v = (0,1,-2)$.
These points satisfy the equation of the plane so they are in the plane. I'll use them as my basis. There are lots of other bases. Any basis will work.

Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -2 & -2 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$ and $A^Tq = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$.

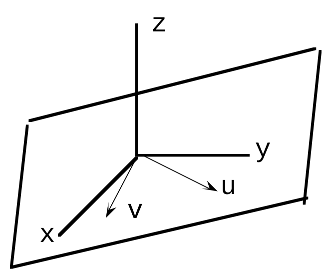
The normal equations are $\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$.

Solution is $a = -2/3$, $b = 4/3$.
so $q_{proj} = -\frac{2}{3}u + \frac{4}{3}v = (-\frac{2}{3}, \frac{4}{3}, -\frac{4}{3})$

$$q - q_{proj} = (\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$$

$$\| q - q_{proj} \| = 1 .$$

So the distance from the point to the plane is 1.



SOLUTIONS Section 3.3

1. (a) The system is

$$\begin{aligned} 2x + y &= 1 \\ x &= 0 \\ x + y &= 0 \end{aligned}$$

The second equation says x = 0. Then the first equation makes y = 1 but the third says y = 0, impossible.

(b) The system

$$\begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

can be rewritten as

$$x \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Since the system has no solution, (1,0,0) is not a combination of (2,1,1) and (1,0,1).

Answer The vector (1,0,0) is not in the subspace spanned by (2,1,1) and (1,0,1).

(c) $A^TA = \begin{bmatrix} 6 & 3 \\ 3 & 2 \end{bmatrix}$, $A^Tb = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

The normal equations are

$$\begin{bmatrix} 6 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\begin{aligned} 6x + 3y &= 2 \\ 3x + 2y &= 1 \end{aligned}$$

Sol is x = 1/3, y = 0.

(d) Part (c) found the projection of $\vec{b}$ onto the subspace of R^3 spanned by u = (2,1,1) and v = (1,0,1).

The projection turned out to be $\frac{1}{3}u + 0v$.

Why use a projection at all? Solving the original system of equations means finding x,y so that xu+ yv *equals* (1,0,0). Since there are no such scalars, the next best thing is to find x,y so that xu + yv is as close to (1,0,0) as possible. That's what the projection does.

(e) sum of square errors = $(2x+y-1)^2 + (x-0)^2 + (x+y-0)^2$.
Its minimum value occurs when x=1/3, y=0. The min value is 1/3.

2. Let $A = \begin{bmatrix} 1 & 1 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 0 \\ 0 & 0 & 3 \\ 1 & 1 & 0 \end{bmatrix}$, $q = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \\ 2 \end{bmatrix}$.

Then $A^TA = \begin{bmatrix} 7 & 2 & -3 \\ 2 & 2 & 1 \\ -3 & 1 & 14 \end{bmatrix}$, $A^Tq = \begin{bmatrix} -3 \\ 2 \\ 4 \end{bmatrix}$.

The normal equations are $A^TA \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^Tq$, i.e.,

$$\begin{aligned} 7x + 2y - 3z &= -3 \\ 2x + 2y + z &= 2 \\ -3x + y + 14z &= 4 \end{aligned}$$

I used a computer to get $x = -\frac{111}{103}$, $y = \frac{219}{103}$, $z = -\frac{10}{103}$.

3 (a) Let the line be $y = mx + b$. Plug in the data points in an attempt to find m and b:

$$\begin{aligned} -m + b &= 2 \\ 0m + b &= 0 \\ m + b &= -3 \\ 2m + b &= -5 \end{aligned}$$

It has to be an inconsistent system of equations because you can see that the points are not collinear (the slope of CD is different from the slope of DE for instance).

In matrix form, the system is

$$A \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -3 \\ -5 \end{bmatrix} \quad \text{where } A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$$

Use the normal equations to get the next best thing to a solution.

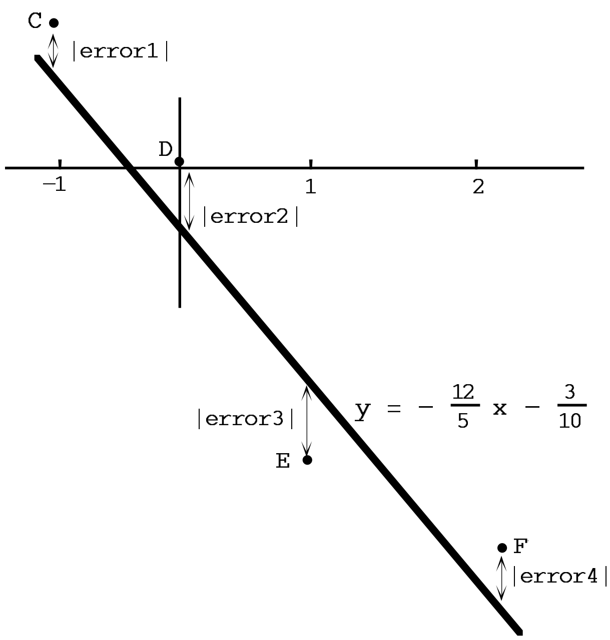
$$A^T A = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix}, \quad A^T \begin{bmatrix} 2 \\ 0 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} -15 \\ -6 \end{bmatrix}$$

Normal equations are

$$\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} -15 \\ -6 \end{bmatrix}$$

Solution is $m = -\frac{12}{5}$, $b = -\frac{3}{10}$. The line of best fit is $y = -\frac{12}{5}x - \frac{3}{10}$.

(b)



(b) In general, for any line $y=mx+b$ that you try to fit to these four data points, the errors are

$$\begin{aligned} \text{error1} &= -m+b-2 \\ \text{error2} &= 0m+b-0 \\ (*) \qquad \text{error3} &= m+b+3 \\ \text{error4} &= 2m+b+5 \end{aligned}$$

For the line of best fit, where $m - -12/5$ and $b = -3/10$, the error specifically are

$$\text{error1} = \frac{12}{5} - \frac{3}{10} - 2 = \frac{1}{10}$$

error2 $= -3/10$

error3 $= 3/10$

error4 = −1/10

(When an error is negative, the point lies above the line because my errors are y-on-line minus y-of-data-point).

(d) Look at all the lines y = mx+b that you could draw. For each line there are the errors given in (*) in part (b).

The line of best fit is the one for which
$(\text{error1})^2 + (\text{error2})^2 + (\text{error3})^2 + (\text{error4})^2$ is minimum.
In this problem the sum of square errors is min using line $y = -\frac{12}{5}x - \frac{3}{10}$. That min value is $\frac{1}{100} + \frac{9}{100} + \frac{9}{100} + \frac{1}{100} = \frac{1}{5}$

4.(a) Need a,b,c so that

$$\begin{aligned} 0a + 0b + c &= 3 \\ a + b + c &= 2 \\ 4a + 2b + c &= 4 \\ 9a + 3b + c &= 4 \end{aligned}$$

The system is inconsistent (there's no way to pass a parabola through the data points). The matrix form of the system is

$$\underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix}}_{A} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \underbrace{\begin{bmatrix} 3 \\ 2 \\ 4 \\ 4 \end{bmatrix}}_{b}$$

Then $A^T b = \begin{bmatrix} 54 \\ 22 \\ 13 \end{bmatrix}$, $A^T A = \begin{bmatrix} 98 & 36 & 14 \\ 36 & 14 & 6 \\ 14 & 6 & 4 \end{bmatrix}$

The normal equations are

$$\begin{bmatrix} 98 & 36 & 14 \\ 36 & 14 & 6 \\ 14 & 6 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 54 \\ 22 \\ 13 \end{bmatrix}$$

The solution is $a = \frac{1}{4}$, $b = -\frac{1}{4}$, $c = \frac{11}{4}$.

The best fitting parabola is $y = \frac{1}{4}x^2 - \frac{1}{4}x + \frac{11}{4}$.

(b) Part (a) found the projection of (3,2,4,4) into the subspace of R^4 spanned by the three cols of A. The projection is $\frac{1}{4}\,\text{col}\,1 - \frac{1}{4}\,\text{col}\,2 + \frac{11}{4}\,\text{col}\,3$.

This is $(\frac{11}{4}, \frac{11}{4}, \frac{13}{4}, \frac{17}{4})$ w.r.t. the usual coord system in R^4 but only the coeffs 1/4, −1/4 and 11/4 were needed in part (a).

5. (a) Need a,b so that

$$\begin{aligned} 2a - 2b &= 0 \\ a + 3b &= 5 \\ 3a + b &= 3 \\ -a + b &- -1 \\ a - b &= 1 \end{aligned}$$

These equations are inconsistent. In matrix form the equations are

$$A \begin{bmatrix} a \\ b \end{bmatrix} = \vec{w} \quad \text{where } A = \begin{bmatrix} 2 & -2 \\ 1 & 3 \\ 3 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \text{ and } w = \begin{bmatrix} 0 \\ 5 \\ 3 \\ -1 \\ 1 \end{bmatrix}$$

Note that the cols of A are orthogonal (call them u and v) so you can avoid the normal equations if you like. Let w_{proj} be the projection of w onto the subspace with orthogonal basis u,v. Then

$$w_{proj} = au + bv \text{ where } a = \frac{u \cdot w}{u \cdot u} = 1,\ b = \frac{v \cdot w}{v \cdot v} = 1.$$

So the least squares solution to the inconsistent system is a = 1, b = 1. And the formula of best fit is z = x + y.

(b) Sum of square errors is

$$(2a-2b-0)^2 + (a+3b-5)^2 + (3a+b-3)^2 + (-a+b+1)^2 + (a-b-1)^2$$

Plug in a=1,b=1 to get the min value 4.

6. It means that when you look at the norm $\|p-v\|$ for each p in the bunch, it's smallest when p is u.

SOLUTIONS Section 3.4

1.(a) orthog basis $u_1 = (0,1,0)$, $u_2 = (1,2,3) - \frac{2}{1}(0,1,0) = (1,0,3)$

orthonormal basis $(0,1,0)$, $\left(\frac{1}{\sqrt{10}}, 0, \frac{3}{\sqrt{10}}\right)$

(b) orthog basis $u_1 = p = (1,2,2)$, $u_2 = (1,3,1) - \frac{9}{9}(1,2,2) = (0,1,-1)$

orthonormal basis $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$, $\left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$

(c) An orthog basis is

$u_1 = (1,0,0,1)$

$u_2 = (1,0,1,0) - \frac{1}{2}(1,0,0,1) = \left(\frac{1}{2}, 0, 1, -\frac{1}{2}\right)$

$u_3 = (2,1,1,1) - \frac{3}{2}(1,0,0,1) - \frac{3/2}{3/2}\left(\frac{1}{2}, 0, 1, -\frac{1}{2}\right) = (0,1,0,0)$

An orthonormal basis is $\left(\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}}\right)$, $\left(\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right)$, $(0,1,0,0)$

2. (a) This question is the same as asking for an orthogonal basis for the 2-dim subspace of R^{34} with basis u and v.
First choose u1 and u2 with the Gram Schmidt formulas using u and v as the x1 and x2.

$u1 = u$

$u2 = v - \frac{v \cdot u1}{u1 \cdot u1}\, u1 = v - \frac{v \cdot u}{u \cdot u}\, u = v - \frac{2}{4}\, u = v - \frac{1}{2} u$

The vectors u1 and u2 are orthogonal.

Check: $u1 \cdot u2 = u \cdot (v - \frac{1}{2}u) = u \cdot v - \frac{1}{2} u \cdot u = 2 - \frac{1}{2}(4) = 0.$

(b) Normalize u1 and u2.

$$u1_{unit} = \frac{u}{\|u\|} = \frac{u}{\sqrt{u \cdot u}} = \frac{1}{2} u$$

$$u2_{unit} = \frac{v - \frac{1}{2}u}{\|v - \frac{1}{2}u\|}$$

where $\|v - \frac{1}{2}u\| = \sqrt{(v - \frac{1}{2}u) \cdot (v - \frac{1}{2}u)} = \sqrt{v \cdot v - u \cdot v + \frac{1}{4} u \cdot u}$

$= \sqrt{3 - 2 + 1} = \sqrt{2}$

So

$$u2_{unit} = \frac{1}{\sqrt{2}} v - \frac{1}{2\sqrt{2}} u$$

3. From #1(c), an orthonormal basis for the subspace is

$$r = \left(\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}}\right),\ s = \left(\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right),\ t = (0,1,0,0)$$

This is an easier basis to use than the original u,v,w because there is a formula for projecting using an orthonormal basis:

$$q_{proj} = (r \cdot q)\ r + (s \cdot q)\ s + (t \cdot q)\ t$$
$$= \frac{10}{\sqrt{2}} r + \frac{12}{\sqrt{2}} s + 6t$$

In terms of the standard basis i,j,k,ℓ for R^4,

$$q_{proj} = \frac{10}{\sqrt{2}}\left(\frac{1}{\sqrt{2}},\ 0,\ 0,\ \frac{1}{\sqrt{2}}\right) + \frac{12}{\sqrt{2}}\left(\frac{1}{\sqrt{6}},\ 0,\ \frac{2}{\sqrt{6}},\ -\frac{1}{\sqrt{6}}\right) + 6(0,1,0,0)$$

$$= (5+\sqrt{12},\ 6,\ 2\sqrt{12},\ 5-\sqrt{12})$$

4. You will end up with an orthogonal basis for R^4. But why bother when you already know a great orthogonal (actually orthonormal) basis for R^4, namely i,j,k,ℓ.

Note that the one you get using the Gram Schmidt process on u,v,w,p (in that order) will not be i,j,k,ℓ. It will include u itself as the first new basis vector.

The only reason I can think of for using the GS process on u,v,w,p is if you want to include one of u,v,w,p in the new basis. For instance if you want a new orthogonal basis that includes w then use the GS process on w,u,v,p (in that order).

It's different if u,v,w,p is a basis say for a 4-dim subspace of R^{27}. Then it is useful to Gram Schmidt them.

SOLUTIONS Section 4.1

1. (a) Rank is 1 (there is maximally one ind col).
(b) Rank is 1 (there is maximally one ind col).
(c) Echelon form (unreduced) is

$$\begin{matrix} 1 & 1 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$$

There are two cols with pivots (cols 1 and 2). Rank is 2.
(d) Maybe you can tell by inspection that the rows of the original matrix are ind. That makes the rank 4.
Or switch rows 3 and 4 to get this unreduced echelon form.

$$\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 1 \end{matrix}$$

Four cols with pivots. So rank is 4.

2. (a) Must be dep since there are maximally 5 ind cols.
(b) Can't tell. You know that some set of 5 cols is ind but you don't know which set.
(c) Can't tell. You know that some set of 5 cols is ind and so any four of those 5 are also ind but you can't tell where any of these cols are.

3. (a) Rank is 7,6,5,4,3,2,1 or 0 (0 only if the matrix is the zero matrix).
(b) The seven rows are dep. But there could be a set of six rows (not containing rows 3,4,7) that are ind. All you can say is that the rank is either 6,5,4,3,2,1 or 0.
(c) Rank is 7 or 6 or 5. Can't be less because the maximal number of ind rows can't be less than 5 since the first five area already independent.
(d) Every set of 6 rows is dep, the set of all 7 rows is dep also since it contains a dependent set (many of them actually) of 6 rows. But there is an ind fivesome of rows. So rank is 5.
(e) Rank is 7 (can't be larger because there are only 7 rows).
(f) The maximal number of ind rows is ≥ 2 (because rows 1 and 2 are ind). So the rank is not 0 or 1.
The rank is not 7 because the set of all 7 rows is dep (since rows 1,2,3 are dep).
Nothing else is ruled out. For instance, there could a set of six ind rows (not a set including rows 1–3 but maybe rows 1,3,4,5,6,7 are ind).
So the rank is 2,3,4,5 or 6.

4. $|M| = 0$ because the echelon form of M is not I (see invertible list).
Rank of M is 2 (because of the 2 echelon cols with pivots).
So some 2×2 subdet is nonzero but you can't tell where it is.

5. The max number of ind cols in M is 5. So the max number of ind rows in M^T is 5. So rank $M^T = 5$.

6. A must be 6×6 . So A^{-1} is 6×6 and A^{-1} is invertible (its inverse is A). The rank of an $n \times n$ invertible matrix is n so the rank of A^{-1} is 6.

7. Tripling all the rows doesn't change the max number of ind rows.
So rank(3B) = 2.

8. *method 1* Use row ops

$R1 \leftrightarrow R2$

$R2 = -\frac{t}{\sqrt{2}} R1 + R2$

$R2 \leftrightarrow R3$

$R3 = \frac{t^2-2}{2} R2 + R3$

to get the unreduced echelon form

$$\begin{array}{ccc} \sqrt{2} & t & \sqrt{2} \\ 0 & \sqrt{2} & t \\ 0 & 0 & \frac{1}{2}t(t^2-4) \\ 2 & & \end{array}$$

I didn't start with the row op $R2 = \frac{-\sqrt{2}}{t} R1 + R2$ because this row op can only be done if $t \neq 0$. So I would have to consider $t = 0$ as a separate case. That's not wrong. But it's a nuisance. That's why I started with $R1 \leftrightarrow R2$.

If $\frac{1}{2}t(t^2-4) \neq 0$ then the three echelon cols have pivots and the rank is 3. This happens when $t \neq 0, \pm 2$.
If $t = 0, \pm 2$ then there are only two echelon cols with pivots and the rank is 2.

method 2

$|M| = t(t^2 - 4)$

$|M| \neq 0$ if $t \neq 0,2,-2$ and in this case the matrix is invertible and the rank is 3.

When $t = 0,2,-2$, the matrix is not invertible so the rank must be lower than 3.
Here is what M looks like when $t = 0,2,-2$.

$$M_{t=0} = \begin{bmatrix} 0 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{bmatrix}, \quad M_{t=2} = \begin{bmatrix} 2 & \sqrt{2} & 0 \\ \sqrt{2} & 2 & \sqrt{2} \\ 0 & \sqrt{2} & 2 \end{bmatrix}, \quad M_{t=-2} = \begin{bmatrix} -2 & \sqrt{2} & 0 \\ \sqrt{2} & -2 & \sqrt{2} \\ 0 & \sqrt{2} & -2 \end{bmatrix}$$

Each matrix has a set of 2 ind cols. So in each of these cases the rank is 2.

9. (a) Rank is 3, 4 or 5.
Can't be 2, 1 or 1 since in that case all 3×3 subdets must be 0.
(b) Same answer as (a).
(c) The whole 5×5 det must be 0 since it is a sum of 4×4 subdets.
All the 4×4 subdets are 0 by hypothesis.
But there is a 3×3 subdet that is not 0.
So the rank is 3.
(d) All the 4×4 subdets must be 0 also (each one expands into a sum of 3×3's). And similarly the whole 5×5 det must be 0. So the rank is 0, 1 or 2 (but not 3,4,5).
(e) Can't conclude anything about rank except that it's ≤ 5 simply because A is 5×5.

SOLUTIONS Section 4.2

1. (a) A no (second col is not normalized)
B yes
reason 1 The rows are orthonormal.
reason 2 The cols are orthonormal.
reason 3 $BB^T = I$.
C no (cols neither orthog nor normalized)
D no (cols not normalized)
E yes *reason 1* The rows are orthonormal:
$$\text{row } 1 \cdot \text{row } 2 = \cos\theta \sin\theta + \sin(-\cos\theta) = 0$$
$$\|\text{row } 1\|^2 = \|\text{row } 2\|^2 = \cos^2\theta + \sin^2\theta = 1$$

reason 2 The cols are orthonormal.
reason 3 $EE^T = I$

2. (a) Yes. And besides, you *do* know something about its columns. If the rows are orthonormal then the cols are automatically orthonormal too.
(b) This isn't possible. If the rows are orthonormal the cols will also be orthonormal.
(c) Not necessarily. Orthogonal is is not good enough. The rows must be orthogonal *and* unit length for the matrix to be orthogonal.

3. (a) No values will do it. Column 2 is not a unit vector. There is no way to make the columns ortho*normal*.

(b) Note that col 2 has unit length so you're not stuck as in part (a).
You need col $1 \cdot$ col $2 = 0$ and $\|\text{col } 1\| = 1$. So you need

$$\frac{a}{\sqrt{10}} + \frac{3b}{\sqrt{10}} = 0$$
$$\sqrt{a^2 + b^2} = 1.$$

which simplifies to

$$a+3b = = 0, \quad a^2 + b^2 = 1$$

Substitute in the second equation for a: $9b^2 + b^2 = 1$, $b = \pm\sqrt{\frac{1}{10}}$

If $b = \sqrt{\frac{1}{10}}$ then $a = -3\sqrt{\frac{1}{10}}$.
If $b = -\sqrt{\frac{1}{10}}$ then $a = 3\sqrt{\frac{1}{10}}$.
So there are two sets of a,b values that work.

4. (a) As a counterexample let
$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$
Then
$$AA^T = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 13 & \cdot \\ \cdot & \cdot \end{bmatrix}$$
but
$$A^TA = \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 20 & \cdot \\ \cdot & \cdot \end{bmatrix}$$
So it can't be true in general that $AA^T = A^TA$.
(b) Let A be orthog. Then $AA^T = I$ and $A^TA = I$. So $AA^T = A^TA$.

5. Diagonal entries must be ±1's. There are 8 such matrices, e.g.,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{etc.}$$

6. (a) Doesn't violate the orthog matrix rule. The rule doesn't apply because the rows are not ortho*normal*.

(c) The matrix B would have orthonormal rows so I expect it to also have orthonormal cols.

Check: $B = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 6/\sqrt{45} & -3/\sqrt{45} \end{bmatrix}$

Norm of col 1 = $\sqrt{\frac{1}{5} + \frac{36}{45}} = \sqrt{\frac{45}{45}} = 1$. Similarly for norm of col 2.

And col 1·col 2 = $\frac{2}{5} - \frac{18}{35} = 0$.

7. (a) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

Then A and B are orthogonal but $A + B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ which is not orthogonal.

(b) Let A and B be orthog. I'll show that $(AB)(AB)^T = I$.

$$\begin{aligned} (AB)(AB)^T &= AB\,B^TA^T && \text{T rule} \\ &= AA^T && BB^T = I \text{ since B is orthog} \\ &= I && AA^T = I \text{ since A is orthog} \end{aligned}$$

(c) If M is orthog then the cols of M are orthonormal. So the rows of M^T are orthonormal So M^T is orthogonal.

(d) Let M be orthogonal. To show that M^{-1} is also orthog I'll show that $M^{-1}(M^{-1})^T$ is I.

version 1

$$\begin{aligned} M^{-1}(M^{-1})^T &= M^{-1}(M^T)^{-1} && \text{by inverse rule} \\ &= M^T\,(M^T)^{-1} && \text{can replace } M^{-1} \text{ by } M^T \text{ since M is orthog} \\ &= I && \text{because the matrix } M^T \text{ times its inverse must be I} \end{aligned}$$

version 2

$$\begin{aligned} M^{-1}(M^{-1})^T &= M^{-1}(M^T)^{-1} && \text{by inverse rule} \\ &= M^{-1}\,(M^{-1})^{-1} && \text{can replace } M^T \text{ by } M^{-1} \text{ since M is orthog} \\ &= I && \text{because the matrix } M^{-1} \text{ times its inverse must be I} \end{aligned}$$

8. No.

reason 1 The cols (and rows) no longer have unit length. Since $\|5u\| = 5\|u\|$, the new cols (and rows) have norm 5, not 1.

reason 2 Let M be $n \times n$. Then $|M| = \pm 1$ so $|5M| = 5^n|M|$ so $|5M| \neq \pm 1$ so 5M can't be orthogonal.

9. By inspection, the rows (cols) of M are orthonormal. So

$$M^{-1} = M^T = \begin{bmatrix} 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

10. Let $P = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_2 & w_3 \end{bmatrix}$. Then $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ and $\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = P^{-1} \begin{bmatrix} x \\ y \\ x \end{bmatrix}$

(a) $x = u_1X + v_1Y + w_1Z$
$y = u_2X + v_2Y + w_2Z$
$z = u_3X + v_3Y + w_3Z$

(b) P is an orthogonal matrix since its columns are orthonormal. So $P^{-1} = P^T$ and

$X = u_1x + u_2Y + u_3Z$
$Y = v_1X + v_2Y + v_3Z$
$Z = w_1X + w_2Y + w_3Z$

11. (a) u,v,w are not ortho*normal* vectors so M is not an orthogonal matrix. But $\frac{1}{2}$ M is orthogonal since its cols are the orthogonal *unit* vectors $\frac{1}{2}$ u, $\frac{1}{2}$ v, $\frac{1}{2}$ w

So $(\frac{1}{2} M)^{-1} = (\frac{1}{2} M)^T$
$2M^{-1} = \frac{1}{2} M^T$ (inverse rule and T rule)
$M^{-1} = \frac{1}{4} M^T$

(b) $\frac{1}{2}$ M is orthogonal so $|\frac{1}{2} M| = \pm 1$, $(\frac{1}{2})^3 |M| = \pm 1$ (det rules), $|M| = \pm 8$.

12. Do these col ops to M:
Divide col 1 by 2
Divide col 2 by 3
Divide col 3 by 4.

Call the result B. Then B is an orthogonal matrix so $|B| = \pm 1$.
By row op rules for dets, $|B| = \frac{1}{2}\frac{1}{3}\frac{1}{4} |M|$.
So $|M| = \pm 24$.

13. The rule says that if a matrix is *square* then if the cols are orthonormal so are the rows. But M isn't square so the rule doesn't apply. In fact the rows of a 3×2 matrix can't be orthogonal because then the rows would be 3 ind vectors in R^2 which is impossible.

SOLUTIONS Review problems for Chapters 3,4

1. Use the Gram Schmidt formulas to get an orthogonal basis u,v first.

Let u = p

$$v = q - \frac{u \cdot q}{u \cdot u} u = \left(\frac{13}{25}, -\frac{6}{25}, \frac{1}{25}, -\frac{12}{25}, 1\right)$$

$\|u\| = 5$, $\|v\| = \frac{1}{5}\sqrt{39}$ so an orthonormal basis is

$$u_{unit} = \left(\frac{2}{5}, \frac{1}{5}, \frac{4}{5}, \frac{2}{5}, 0\right), \quad v_{unit} = \frac{5}{\sqrt{39}} v.$$

2. You need k so that the three vectors are ind.

method 1 (easiest)

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & k \\ 1 & 4 & k^2 \end{vmatrix} = k^2 - 3k + 2 = (k-2)(k-1)$$

To get ind cols you want a nonzero det so you need $k \neq 2$ and $k \neq 1$.

So any value of k is OK except 1 and 2.

method 2 Use row ops:

$$\begin{matrix} 1 & 1 & 1 \\ 1 & 2 & k \\ 1 & 4 & k^2 \end{matrix} \sim \begin{matrix} 1 & 1 & 1 \\ 0 & 1 & k-1 \\ 0 & 3 & k^2-1 \end{matrix} \sim \begin{matrix} 1 & 1 & 1 \\ 0 & 1 & k-1 \\ 0 & 0 & -3(k-1)+k^2-1 \end{matrix}$$

To get ind cols you need $-3(k-1) + k^2 - 1 \neq 0$, $k^2 - 3k + 2 \neq 0$, $k \neq 2,1$.

3. Want to show that $(B^{-1}SB)^T = B^{-1}SB$

$(B^{-1}SB)^T = B^T S^T (B^{-1})^T$ T rule

$= B^{-1} S (B^T)^T$ since $B^{-1} = B^T$ and $S^T = S$

$= B^{-1}SB$ QED

4. (a) Line up all the vectors as cols and row op.

$$\text{old} = \begin{matrix} 1 & 2 & 0 & a \\ 0 & 1 & 1 & b \\ 2 & 3 & 1 & c \end{matrix} \qquad \text{echelon} = \begin{matrix} 1 & 0 & 0 & -a-b+c \\ 0 & 1 & 0 & a+\frac{1}{2}b-\frac{1}{2}c \\ 0 & 0 & 1 & -a+\frac{1}{2}b+\frac{1}{2}c \end{matrix}$$

The first 3 echelon cols have pivots so u,v,w are ind. So u,v,w are a basis for R^3 (any 3 ind vectors in R^3 are a basis).

In the echelon form, $C_4 = (-a-b+c)C_1 + (a + \frac{1}{2}b - \frac{1}{2}c)C_2 + (-a + \frac{1}{2}b + \frac{1}{2}c)C_3$.

(b) Let

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

$|P| \neq 0$ so the cols u,v,w are ind. So u,v,w is a basis for R^3 (any 3 ind vectors in R^3 are a basis).

$$P^{-1} = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 1/2 & -1/2 \\ -1 & 1/2 & 1/2 \end{bmatrix}, \qquad P^{-1}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -a-b+c \\ a + \frac{1}{2}b - \frac{1}{2}c \\ -a + \frac{1}{2}b + \frac{1}{2}c \end{bmatrix}$$

The new coords of (a,b,c) are $-a-b+c$, $a + \frac{1}{2}b - \frac{1}{2}c$, $-a + \frac{1}{2}b + \frac{1}{2}c$

So $(a,b,c) = (-a-b+c)u + (a + \frac{1}{2}b - \frac{1}{2}c)v + (-a + \frac{1}{2}b + \frac{1}{2}c)w$.

5. I'll show that $\|Mu\|^2 = \|u\|^2$.

$\|Mu\|^2$	$= Mu \cdot Mu$	norm property
	$= (Mu)^T Mu$	connection between dots and matrix mult
	$= u^T M^T Mu$	T rule
	$= u^T u$	M is orthog so $M^T M$ is I
	$= u \cdot u$	connection between dots and matrix mult
	$= \|u\|^2$	

So $\|Mu\| = \|u\|$. QED

6. (a) Look at the echelon form where u,v,w,q were the original cols.
Cols 1 and 3 have pivots. So a basis for the subspace is u,w.
The subspace is 2-dim so any two ind vectors in the subspace are a basis.
Echelon cols 2 and 3 are ind so another basis is v,w. Also w,q etc.
The only pair that isn't a basis is u,v.
(b) Look at the echelon form where u,v,w,q were the original rows.
The non-zero echelon rows r = (1,1,0,0,0,0), s = (0,0,1,1,1,1) are the standard basis for the subspace.
(c) The subspace consists of all vectors of the form

$$a(1,1,0,0,0,0) + b(0,0,1,1,1,1) = (a,a,b,b,b,b)$$

So the subspace consists of all vectors whose first two components are equal and whose last three components are equal.
For example, some vectors in the subspace are (2,2,3,3,3,3), (e,e,π,π,π,π), (5,5,5,5,5), etc.
(d) Since the subspace is 2-dim, any two ind vectors in the subspace are a basis.

basis 1 (2,2,9,9,9,9), (π,π,e,e,e,e)
basis 2 (6,6,6,6,6,6) (1,1,−1,−1,−1,−1) etc.

7. Let the line be y= mx + b. Then you want m and b so that

$$0 = m + b$$
$$2 = 2m + b$$
$$7 = 3m + b$$

This is an inconsistent system. Get the least squares solution instead. The system of equations is

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}}_{A} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 7 \end{bmatrix}$$

Switch to the system of equations

$$A^T A \begin{bmatrix} m \\ b \end{bmatrix} = A^T A \begin{bmatrix} 0 \\ 2 \\ 7 \end{bmatrix}$$

$$\begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 25 \\ 9 \end{bmatrix}$$

Solution is m = 7/2, b = −4.
So the line of best fit is $y = \frac{7}{2} x - 4$.

8. (a) *method 1* Line up u,v,w,p as columns and find the echelon form.
If all the echelon columns have pivots then u,v,w,p are ind. Otherwise they are dep.

method 2 Line up u,v,w,p as rows and find the echelon form. If all the echelon rows are nonzero then u,v,w,p are ind. Otherwise they are dep.

(b) Line up u,v,w,y as columns (in that order) and find the echelon form.

case 1 The last echelon col has a pivot.

Then the last echelon col is *not* a comb of the preceding echelon cols and in that case y is not a comb of u,v,w so it is not in the subspace spanned by u,v,w.

case 2 The last echelon col doesn't have a pivot.

Then the last echelon col *is* a comb of the preceding echelon cols and in that case y is a comb of u,v,w so it is in the subspace spanned by u,v,w.

(c) If the echelon form of A is I then A is invertible. If it's not I then A is not invertible.

(d) Count the number of echelon cols with pivots. Or count the number of nonzero echelon rows.

SOLUTIONS Section 5.1

1. (a) $x_6 = 7$
$x_3 = 6 - 6x_5 - 4x_4$
$x_1 = 5 - 5x_5 - 3x_4 - 2x_2$
x_2, x_4, x_5 free

The solution can also be written as
$x_6 = 7$
$x_5 = t$
$x_4 = s$
$x_3 = 6 - 6t - 4s$
$x_2 = r$
$x_1 = 5 - 5t - 3s - 2r$

(b) No sols because of row 3

(c) $x_1 = 5$, $x_2 = 6$ (no free variables)

(d) $x_1 = 0$; x_2, x_3, x_4 free

The solution can also be written as $x_1 = 0$, $x_2 = r$, $x_3 = s$, $x_4 = t$.

(e) $x_1 = 5 - 2x_2$
$x_3 = 3$
x_2, x_4 free

The solution can also be written as
$x_1 = 5 - 2t$
$x_2 = t$
$x_3 = 3$
$x_4 = s$

2. (a) *method 1* Line up u,v,w,y as cols:

$$\begin{matrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{matrix}$$

Do the row ops
row3 = −row1 + row3
row4 = −row1 + row4;
row3 = −row2 + row3;
to get

$$\begin{matrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}$$

You can see that in the echelon form, $C_4 = 2C_1 - C_2$ so y is in the subspace spanned by u,v,w and in particular y = 2u − v.
There are other ways to write y as a comb of u,v,w. In the echelon form,
$C_4 = 3C_1 - C_3$ so y = 3u − w etc.
The subspace is 2-dim because the first three echelon cols are dep and and the first two are ind.

method 2 Line up u,v,w (not y) as rows and row op into echelon form:

$$\begin{matrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 2 & 1 \end{matrix} \quad \text{row ops to} \quad \begin{matrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}$$

The nonzero echelon rows are a basis for the echelon ow space and for the original

row space as wll. So the space spanned by u,v,w is the set of vectors of the form a(1,0,1,1) + b(0,1,1,0), i..e., of the form (a,b, a+b, a).
By inspection, y is of this form (with a = 2, b = −1) so y is in the subspace spanned by u,v,w,.

(b) Solve y = au + bv + cw

$$\begin{aligned} a + c &= 2 \\ b + c &= -1 \\ a + b + 2c &= 1 \\ a + c &= 2 \end{aligned}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{array}\right]$$

Do the row ops from part (a) to get

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Sol is

$$\begin{aligned} a &= 2-c \\ b &= -1 - c \\ c &= \text{anything} \end{aligned}$$

Since the system of equations is consistent, y is in the subspace spanned by u,v,w.
Since we got more than one solution, the vectors u,v,w can't be a basis (if u,v,w were a basis then the coords of y w.r.t. u,v,w would be unique and there would have been just one solution).
To express y in terms of u,v,w choose say c = 0. Then a = 2, b = −1 and y = 2u − v.
Or you could choose c = 28. then a = −26, b = −29 and y = −26u − 29v + 28w etc.

3. Start with

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 2 & k & 6 & 6 \\ -1 & 3 & k-3 & 0 \end{array}\right]$$

and do row ops

$$\begin{aligned} R2 &= -2R1 + R2 \\ R3 &= R1 + R3 \\ R3 &\leftrightarrow R2 \\ R3 &= -(4+k)R2 + R3 \end{aligned}$$

to get

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 0 & 1 & k & 1 \\ 0 & 0 & -k(k+4) & -k \\ 0 & 0 & 0 & 0 \end{array}\right]$$

case 1 k = 0 Then the system is

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and it's consistent. There is a col without a pivot (the third) so there are infinitely many solutions.
case 2 k = −4 Then the system is

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and is inconsistent.
case 3 k ≠ 0,4 Then the system is

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 0 & 1 & k & 1 \\ 0 & 0 & \text{nonzero} & -k \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and there is one solution.

4. (a) False.

There can be no sols, e.g., $\left[\begin{array}{cccc|c} 1 & 0 & 2 & 4 & 5 \\ 0 & 1 & 3 & 5 & 6 \\ 0 & 0 & 0 & 0 & 7 \end{array}\right]$

There can be infinitely many solutions, e.g., $\left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 5 \\ 0 & 1 & 0 & 3 & 6 \\ 0 & 0 & 1 & 0 & 7 \end{array}\right]$

Can't have exactly one solution because it isn't possible for all the echelon cols to have pivots.

(b) False.

System can have no sols, e.g., $\left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 6 \\ 0 & 0 & 7 \end{array}\right]$

Can have one solutions, e.g., $\left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 6 \\ 0 & 0 & 0 \end{array}\right]$

Can have infinitely many, e.g. $\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$

(c) False. Can have one solution, e.g, $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 4 \end{array}\right]$

Can have have no sols, e.g., $\left[\begin{array}{ccc|c} 1 & 0 & 2 & 5 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 7 \end{array}\right]$

Can have infinitely many sols, e.g., $\left[\begin{array}{ccc|c} 1 & 0 & 2 & 5 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right]$

5. Rank A = 2, number of variables = 7 so n−r = 5. Either there are no sols or there are infinitely many with 5 free variables.

6. (a) $x_3 = \dfrac{-2 - 2x_5 - 7x_4}{5}$

$$x_1 = \frac{4 - x_5 - 6x_4 - 4x_3 - 3x_2}{2} = \frac{28}{10} + \frac{3}{10}x_5 - \frac{2}{10}x_4 - \frac{3}{2}x_2$$

x_4, x_5, x_2 free

(b) Use row ops
R3 = − R1 + R3
R3 = −R2 + R3
R4 = −R2 + R4
R3 ↔ R4
to get

$$\left[\begin{array}{ccccc|c} 2 & 8 & 5 & 2 & -6 & 8 \\ 0 & 4 & 6 & 3 & 3 & 9 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Sol is

$v = 0$

$$y = \frac{9 - 3v - 3u - 6z}{4} = \frac{9 - 3u - 6z}{4}$$

$$x = \frac{8 - 2u - 5z - 3y}{2} = -5 + \frac{7}{2}z + 2u$$

u,z free

(c) row2 = −2 row1 + row2
row3 = row1 + row3;
row3 = 3 row2 + row3;

Get

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 0 & -4 & -4 \end{array}\right]$$

Sol is

$z = 1$

$y = 4 - 2z = 2$

$$x = \frac{1 - y - z}{2} = -1$$

(no free variables)

7. (a) $z = 6 - 4w$, $x = 5 - 3w - 2y$

(b) $z = 6 - 4w$

$$y = \frac{1}{2}(5 - 3w - x)$$

(c) $w = \frac{1}{4}(6 - z)$

$$x = 5 - 3w - 2y = 5 - 3 \cdot \frac{1}{4}(6-z) - 2y = \frac{1}{2} + \frac{3}{4}z - 2y$$

(d) Not possible.

8. A|b row ops to

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \cdot \\ 0 & 1 & 0 & \cdot \\ 0 & 0 & 1 & \cdot \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So A|c has either no sols (if the row-opped $\vec{c}$ has a nonzero 4th entry) or one solution (if the row-opped $\vec{c}$ has a zero 4th entry).

9. (a) A has 3 echelon col with pivots; A|b row ops to something like this:

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 4 & 0 & \cdot \\ 0 & 0 & 1 & 5 & 0 & \cdot \\ 0 & 0 & 0 & 0 & 1 & \cdot \end{array}\right]$$

There is no room for an all zero echelon row so it's not possible for Ax = b to be inconsistent. There must be infinitely many solutions with 2 free variables.

(b) A has 2 echelon cols with pivots.

Could have infinitely many sols with 3 free variables. Or could have no sols if the system row ops to something like

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 4 & 6 & \cdot \\ 0 & 0 & 1 & 5 & 7 & \cdot \\ 0 & 0 & 0 & 0 & 0 & \text{nonzero} \end{array}\right]$$

(c) A has 3 echelon cols with pivots so A|b row ops to

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \cdot \\ 0 & 1 & 0 & \cdot \\ 0 & 0 & 1 & \cdot \\ 0 & 0 & 0 & \cdot \\ 0 & 0 & 0 & \cdot \end{array}\right]$$

Could be no sols, if the 4th or 5th components in the row-opped b are nonzero
Otherwise there is one sol.

10. Let $\vec{b} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$. The system row ops to $\left[\begin{array}{ccc|c} 1 & 2 & 0 & a-b \\ 0 & 0 & 1 & b \\ 0 & 0 & 0 & c-2b \\ 0 & 0 & 0 & d-3b \end{array}\right]$

To get a sol you need $c-2b = 0$, $d-3b = 0$.
So any a,b,c,d are OK as long as $d = 3b$, $c = 2b$.
In that case the solution is $z = b$, $x = a - b - 2y$ (y is free).

11. Solving the matrix equation AB = I amounts to solving these three systems (see observation (3) about matrix multiplication in Section 1.2):

$$(*) \qquad A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad A \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad A \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The echelon form of A has 3 echelon cols with pivots (since rank A is 3); the systems in (*) row ops to something like

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 2 & \cdot \\ 0 & 1 & 0 & 3 & \cdot \\ 0 & 0 & 1 & 4 & \cdot \end{array}\right] \quad \text{or} \quad \left[\begin{array}{cccc|c} 1 & 2 & 0 & 0 & \cdot \\ 0 & 0 & 1 & 0 & \cdot \\ 0 & 0 & 0 & 1 & \cdot \end{array}\right] \quad \text{etc.}$$

There can't be a row of 0's in the echelon form of A so each of the three systems in (*) is consistent, i.e., you can solve for all the unknowns.
In fact, each system in (*) not only is consistent but has infinitely many solutions. The echelon form of A has 3 cols with pivots so each system has one free variable. So there are infinitely many B's.

SOLUTIONS Section 5.2

1. No row ops necessary. Just back substitute.

$x_5 = 0$

$x_4 = s$

$x_3 = t$

$x_2 = -3t - 5s$

$x_1 = 4x_4 - 2x_3 - 5x_2 = 12t + 21s$

One way to get a basis is to let t = 0, s = 1 and then let t = 1, s = 0. This gets solutions u = (21,−5,0,1,0), v = (13,−3,1,0,0). A basis is u,v.

(b) System is $\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 3 & 2 & 0 \end{array}\right]$ It row ops to $\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$

Sol is $x_2 = -x_3$, $x_1 = -x_3 - 2x_2 = x_3$.

The solution has parametric equations $x_1 = t$, $x_2 = -t$, $x_3 = t$.

Subspace is 1-dim. To get a basis, just get one (nonzero) sol.
If t = 1 then sol is u = (1,−1,1); u is a basis for the subspace.

(c) System is $\left[\begin{array}{ccc|c} 2 & 1 & 1 & 0 \\ 4 & 1 & 0 & 0 \\ -2 & 2 & 1 & 0 \end{array}\right]$ It row ops to $\left[\begin{array}{ccc|c} 2 & 1 & 1 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & -4 & 0 \end{array}\right]$

Sol is $x_3 = 0$, $x_2 = 0$, $x_1 = 0$.

The subspace of solutions contains only $\vec{0}$. It's a 0-dim subspace and it isn't considered to have a basis.

2. (a) The system $\left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 \end{array}\right]$ row ops to $\left[\begin{array}{cccc|c} 1 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

Sol is $x_3 = -x_4$, $x_1 = 2x_4 - x_2$ (x_2, x_4 free).

Null space of M is 2-dim (two free variables). To find two ind sols to be the basis vectors, set $x_2 = 1$, $x_4 = 0$ and then set $x_2 = 0$, $x_4 = 1$ to get u = (−1,1,0,0), v = (2,0,−1,1).

(b) Solve $\left[\begin{array}{cccc|c} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

Sol is $x_2 = x_3 = x_4 = 0$, x_1 free.

The null space of M is 1-dim with (pick, say, x_1=1) basis vector (1,0,0,0).

(c) The system of equations $M\vec{x} = \vec{0}$ is x + y + 2z = 0. The solution is x = −y−2z; y,z free. Null space is 2-dim.
To find basis vectors choose y=0,z=1 and then use y=1,z=0.
A basis is (−2,0,1), (−1,1,0).

3. System is $\left[\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 0 \\ 2 & 5 & 3 & -1 & 0 \end{array}\right]$ It row ops to $\left[\begin{array}{cccc|c} 1 & 0 & -1 & -3 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{array}\right]$

Sol is $x_2 = -x_3 - x_4$, $x_1 = 3x_4 + x_3$.

To get two ind vectors for the basis, set $x_3 = 1$, $x_4 = 0$ to get p = (1,−1,1,0); set $x_3 = 0$, $x_4 = 1$ to get q = (3,−1,0,1).

To get an orthogonal basis use the Gram Schmidt process on p,q. Take

$\vec{u_1} = p$

$$\vec{u_2} = q - \frac{u_1 \cdot q}{u_1 \cdot u_1} u_1 = \left(\frac{5}{3}, \frac{1}{3}, -\frac{4}{3}, 1\right)$$

Orthog basis is $\vec{u_1}$, $\vec{u_2}$.

4. In each case, the set of solutions must either be a plane through the origin, a line through the origin, just the origin, or all of R^3 since these are the only subspaces of R^3.

(a) $x = 0$; y,z free. The set of solutions is the y,z plane.

(b) y,z free; $x = -\frac{1}{2}(3y + 4z)$. The set of sols is the plane $2x + 3y + 4z = 0$.

(c) x,y,z, free. The set of solutions is all of R^3.

(d) y,z free; $x = -3y$. The set of solutions is plane $x + 3y = 0$.

(e) $x = 2z$, $y = 3z$, z free.
The set of solutions is the line with parametric equations $x = 2t$, $y = 3t$, $z = t$.

5. (a) All vectors in R^5 are solutions.
(b) There are no solutions.

6. For each system, $n = 6$, $r = 5$, $n-r = 1$.
(a) Infinitely many sols with one free variable.
(b) Either no sols or infinitely many sols with one free variable.

7. (a) The cross product $u \times v = (-2,-2,1)$ is perp to u,v.
The only other perps to u and v are multiples of $(-2,-2,1)$.

(b) Let $\vec{x} = (x,y,z)$. Solve the system of equations $\vec{u}\cdot\vec{x} = \vec{0}$, $\vec{v}\cdot\vec{x} = \vec{0}$.

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 1 & 1 & 4 & 0 \end{array}\right] \text{ row ops to } \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$$

Sol is $y = -2z$, $x = -2z$, z free.
The sol is a 1-dim subspace of R^3.
Set $z = 1$ to get $w = (-2,-2,1)$, a basis vector for the subspace.
The set of vectors perp to u and v is the set of multiples of $(-2,-2,1)$.

8. Solve the equation $\vec{u}\cdot(x_1,x_2,x_3,x_4) = 0$, i.e.,

$$(*) \qquad x_1 + x_2 + 4x_3 - 2x_4 = 0.$$

This is a homog system of equations, one equation in 4 unknowns.
One way to write the solution is

$$x_1 = 2x_4 - 4x_3 - x_2; \quad x_2,x_3,x_4 \text{ free.}$$

The solutions are a 3-dim subspace of R^4 so you can get 3 ind vectors orthogonal to u but no more than that. Here is one set of 3 ind solutions (actually a basis for the subspace of solutions to (*)).

Let $x_2 = 1$, $x_3 = 0$, $x_4 = 0$ to get solution p = (−1,1,0,0)
Let $x_2 = 0$, $x_3 = 1$, $x_4 = 0$ to get solution q = (−4,0,1,0)
Let $x_2 = 0$, $x_3 = 0$, $x_4 = 1$ to get solution r = (2,0,0,1).
Three ind solutions are p,q,r.

9. (a) The given equations are a homogeneous system of two equations in 6 unknowns:

$$x_2 + x_3 - x_5 = 0$$
$$2x_3 - x_4 + x_5 = 0$$

(There are six unknowns even though they don't all actually appear in the equations because the problem is about points of the form $(x_1,\ldots, x_6)$.)

So the set of solutions is a subspace of R^6.
The solution is

$$x_2 = -x_3 + x_5$$
$$x_4 = 2x_3 + x_5$$
$$x_1,\ x_3,\ x_5,\ x_6 \text{ free}$$

To get a basis
let $x_1 = 1$, $x_3 = 0$, $x_5 = 0$, $x_6 = 0$ to get sol (1,0,0,0,0,0);
let $x_1 = 0$, $x_3 = 1$, $x_5 = 0$, $x_6 = 0$ to get sol (0,−1,1,2,0,0);
let $x_1 = 0$, $x_3 = 0$, $x_5 = 1$, $x_6 = 0$ to get sol (0,1,0,1,1,0);
let $x_1 = 0$, $x_3 = 0$, $x_5 = 0$, $x_6 = 1$ to get sol (0,0,0,0,0,1).
A basis is (1,0,0,0,0,0), (0,−1,1,2,0,0), (0,1,0,1,1,0), (0,0,0,0,0,1).

(b) This is the set of points $(x_1,\ldots, x_5)$ where

$$x_2 = x_1$$
$$x_3 = 2x_1$$
$$x_4 = 2x_1$$
$$x_5 = \text{anything}$$

It's a homogeneous system of 3 equations in 5 unknowns:

$$x_1 - x_2 = 0$$
$$2x_1 - x_3 = 0$$
$$2x_1 - x_4 = 0$$

The solution has free variables x_1, x_5 so the solutions are a 2-dim subspace of R^5.
To get a basis, first plug in $x_1=1, x_5=0$ and then $x_1=0$, $x_5=1$. A basis is (1,1,2,2,0), (0,0,0,0,1)

10. The set of vectors $\vec{x}$ such that $A\vec{x} = B\vec{x}$ is the set of solutions to the homogeneous system of equations $(A-B)\vec{x} = \vec{0}$. So the set is a subspace of R^7.

11. Solution is y = 0, x free. The space of solutions is 1-dim. A basis vector is i = (1,0).

SOLUTIONS Section 5.3

1. Solve $\left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{array}\right]$ Solution is

$x_2 = -x_4$

$x_1 = -x_2 = x_4$

x_3, x_4 free

If $x_3 = 1$, $x_4 = 0$ then the sol is u = (0,0,1,0).

If $x_3 = 0$, $x_4 = 1$ then the sol is v = (1,−1,0,1).

Orthog complement is a 2-dim subspace with basis u,v.

2. Find the orthog comp of the subspace spanned by p,q,r. Solve

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{array}\right]$$

Row op to

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{array}\right]$$

Sol is $x_3 = -x_4$

$x_2 = x_6$

$x_1 = -x_6$

x_4, x_5, x_6 free

So you can say that the answer is the set of vectors of the form $(x_1,\ldots,x_6)$ where $x_3 = -x_4$, $x_2 = x_6$ and $x_1 = -x_6$. Or you can pick 3 ind sols like u = (0,0,−1,1,0,0), v = (0,0,0,0,1,0), w = (−1,1,0,0,0,1) and say that the answer is the 3-dim subspace with basis u,v,w.

3. The echelon form of A is

$$\begin{array}{ccccc} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}$$

(a) The row space is a 3-dim subspace of R^5 with basis (1,2,0,3,0), (0,0,1,4,0), (0,0,0,0,1).

(b) Col space is a 3-dim subspace of R^4 with basis (1,2,0,1), (0,1,1,1), (0,0,1,0), the original cols corresponding to the echelon cols with pivots.

(c) A has 5 cols and rank 3 so n = 5, r = 3 and dim of null space is n−r = 2.

To find the basis, solve the system

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Solution is

$x_1 = -3x_4 - 2x_2$

$x_3 = -4x_4$

$x_5 = 0$

x_2, x_4 free

If $x_2 = 1$, $x_4 = 0$ then the sol is $u = (-2,1,0,0,0)$.
If $x_2 = 0$, $x_4 = 1$ then the sol is $v = (-3,0,-4,1,0)$.
A basis for the null space is u,v.

(d) The row space of A and the null space of A are orthog complements. So the orthog complement of the null space is the row space. From part (a), the orthog comp is a 3-dim subspace of R^5 with basis (1,2,0,3,0), (0,0,1,4,0), (0,0,0,0,1).

4. (a) First of all, the second col of A is a multiple of the first, so the col space is spanned by cols 1 and 3 alone (or by cols 2 and 3 alone).

To get the orthog complement of the space spanned by cols 1 and 3 of A, line up the spanning vectors as rows of a matrix and find the null space. So let

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

and solve $B\vec{x} = \vec{0}$. By inspection,

$$x_1 = 0,\ x_2 = -\frac{4}{3}x_3 - \frac{5}{3}x_4.$$

Any two ind sols can serve as a basis.
One possibility is is (0,−4,3,0), (0,−5,0,3).

(b) Use the Gram Schmidt process on $\vec{x}_1 = (0,-4,3,0)$, $\vec{x}_2 = (0,-5,0,3)$.
New orthogonal basis is

$$\vec{u}_1 = \vec{x}_1 = (0,-4,3,0)$$

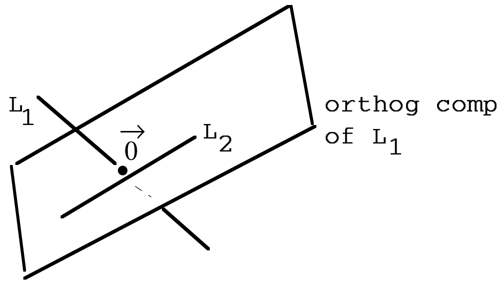
$$\overrightarrow{u_2} = \overrightarrow{x_2} - \frac{u_1 \cdot x_2}{u_1 \cdot u_1}\overrightarrow{u_1} = (0,-5,0,3) - \frac{20}{25}(0,-4,3,0) = \left(0, -\frac{9}{5}, -\frac{12}{5}, 3\right)$$

5. First of all, the hyperplane is a subspace of R^5 because it's the set of solutions to a homogeneous system (one equation in five unknowns). In particular, the coeff matrix A is the 1×5 matrix $\begin{bmatrix} 2 & 3 & 5 & -7 & 1 \end{bmatrix}$. The hyperplane is the null space of A.

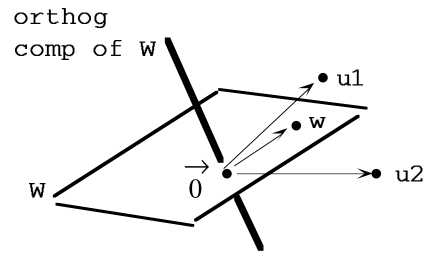
The set of vectors orthog to the hyperplane is the orthog complement of the hyperplane, namely, the row space of A. This row space consists of all multiples of the one row $\vec{n} = (2,3,5,-7,1)$. So $\vec{n}$ and multiples of $\vec{n}$ are the only vectors orthog to the hyperplane. QED

By the way, the hyperplane has dimension 4 because rank A = 1 and n = 5 (or equivalently because when you solve the system of one equation in five unknowns, there are 4 free variables). In general, a hyperplane in R^n has an equation of the form $a_1x_1 + \dots + a_nx_n = 0$ and is an (n−1)-dimensional subspace of R^n. The orthogonal complement is a 1-dimensional subspace (a "line through the origin") with basis vector $(a_1, \dots, a_n)$.

6. No. Line L_2 does not contain *all* vectors orthogonal to all the vectors in L_1. It only contains some of them. It's the plane through the origin perp to L_1 that is the subspace containing all of them. So the orthogonal complement of L_1 is a plane (see the diagram). Similarly, the orthogonal complement of L_2 is the plane through the origin perp to L_2.



Problem 6



Problem 7(b)

7. (a) u is orthog to w because u is orthogonal to *every* vector in W.

(b) Can't tell. Since u is not in the orthogonal complement of W you know that u cannot be orthog to *every* vector in W. In other words, there is at least one vector in W that u is not orthog to. But that "at least one" vector might not be w.

The diagram shows a subspace W in R^3, a plane through the origin, and its orthogonal complement, a line through the origin perp to the plane. The vector w is in W. There are two u vectors, neither in the orthog comp of W. I tried to make the first look like it is not orthog to v and the second look like it is orthog to v.

8. No. the row space of a matrix and its null space are orthogonal complements. so a vector in the null space has to be orthog to everything in the row space. But (2,1,1) is not orthog to (3,1,2).

SOLUTIONS Section 5.4

1. M violates (7) on the invertible list so nothing on the list holds. For instance, M is not invertible.

2. (a) $\begin{vmatrix} 2 & -5 & 4 \\ 1 & -3 & 2 \\ 0 & 1 & 1 \end{vmatrix} = -1$ so there is just one sol.

$$x = \frac{\begin{vmatrix} 5 & -5 & 4 \\ 1 & -3 & 2 \\ 7 & 1 & 1 \end{vmatrix}}{-1} = 2, \qquad y = \frac{\begin{vmatrix} 2 & 5 & 4 \\ 1 & 1 & 2 \\ 0 & 7 & 1 \end{vmatrix}}{-1} = 3 \qquad z = \frac{\begin{vmatrix} 2 & -5 & 5 \\ 1 & -3 & 1 \\ 0 & 1 & 7 \end{vmatrix}}{-1} = 4$$

(b) $\begin{vmatrix} 0 & 2 & 4 \\ 1 & 1 & 4 \\ 1 & 0 & 2 \end{vmatrix} = 0$. System has either no sols or infinitely many. Use row ops to find out which it is. The system row ops to

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 10 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

No solutions because of the last line.

(c) Same coeff matrix as part (b). System row ops to

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Infinitely many sols.

$y = 6 - 2z$
$x = 4 - z$
z free

3. (a) Matrix of coeffs is $M = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}$.

There is exactly one sol because $|M| \neq 0$.

$$\text{Sol is} \quad x = \frac{\begin{vmatrix} 3 & 1 \\ 5 & -1 \end{vmatrix}}{-5} = \frac{8}{5}, \qquad y = \frac{\begin{vmatrix} 3 & 3 \\ 2 & 5 \end{vmatrix}}{-5} = -\frac{9}{5}$$

(b) System is $M\vec{x} = \vec{b}$ where $M = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

Solution is $\vec{x} = M^{-1}\vec{b} = \dfrac{1}{-5}\begin{bmatrix} -1 & -1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 8/5 \\ -9/5 \end{bmatrix}$

(c) $M|b$ row ops to $\left[\begin{array}{cc|c} 1 & 0 & 8/5 \\ 0 & 1 & -9/5 \end{array}\right]$

Sol is $x = \frac{8}{5}$, $y = -\frac{9}{5}$.

4. $$x = \frac{\begin{vmatrix} c & b \\ f & e \end{vmatrix}}{\begin{vmatrix} a & b \\ d & e \end{vmatrix}} = \frac{ec - bf}{ae - bd}, \qquad y = \frac{\begin{vmatrix} a & c \\ d & f \end{vmatrix}}{\begin{vmatrix} a & b \\ d & e \end{vmatrix}} = \frac{af - cd}{ae - bd}$$

assuming $ae - bd \neq 0$, i.e., assuming $\begin{vmatrix} a & b \\ d & e \end{vmatrix} \neq 0$.

5. A can't be invertible (if it were, then Ax=b would have had one solution).

(a) Can't be just one solution because A is not invertible.
Can't be no solutions since a homog system always has at least the trivial solution. So must be infinitely many solutions. There would be either 1 or 2 or 3 free variables (there would be 3 free variables iff A were the zero matrix).
(b) Can't be just one sol because A is not invertible. There are either no solutions or infinitely many.

If A|c row opped to something like $\left[\begin{array}{ccc|c} 1 & 2 & 0 & 5 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 7 \end{array}\right]$ then there would be no solutions.

If A|c row opped to something like $\left[\begin{array}{ccc|c} 1 & 2 & 0 & 5 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right]$ then there would be infinitely many solutions.
(c) $|A| = 0$ since A is not invertible.
(d) Rank can't be 3. Rank must be 2 or 1 or 0 (in the special case that A is the zero matrix).

6. The rows of A are orthonormal (by inspection) so the matrix A is orthogonal. So A is invertible and in particular $A^{-1} = A^T$.

Solution is $$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^T \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{5} & 1/\sqrt{30} \\ 2/\sqrt{6} & 1/\sqrt{5} & 2/\sqrt{30} \\ 1/\sqrt{6} & 0 & 5/\sqrt{30} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

$$x = \frac{1}{\sqrt{6}} - \frac{6}{\sqrt{5}} + \frac{4}{\sqrt{30}}$$

$$y = \frac{4}{\sqrt{6}} + \frac{3}{\sqrt{5}} + \frac{8}{\sqrt{30}}$$

$$z = \frac{2}{\sqrt{6}} + \frac{20}{\sqrt{30}}$$

SOLUTIONS Section 6.1

1. Want a matrix M so that $M\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_4 \\ x_4 \\ 2x_1 - 3x_2 \\ 5x_1 + 6x_3 + 7x_4 \end{bmatrix}$

So $M = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 2 & -3 & 0 & 0 \\ 5 & 0 & 6 & 7 \end{bmatrix}$

2. $M \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$ so the image is (10,10).

To get the pre-image of (0,1) solve $M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Sol is $\begin{bmatrix} x \\ y \end{bmatrix} = M^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 1/5 \end{bmatrix}$

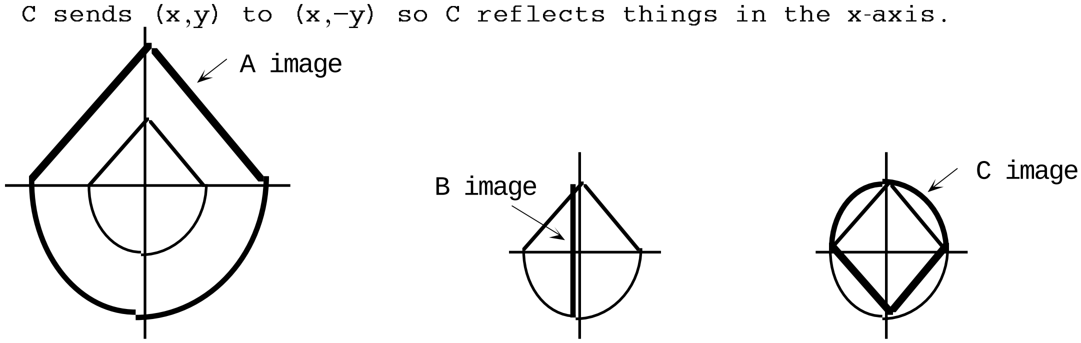
So the pre-image is (−2/5, 1/5).

3. (a) $A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$ so A sends point (x,y) to point (2x,2y).

Every point moves away from the origin so that it is twice as far from the origin as it used to be.

B maps (x,y) to (0,y) so B projects points on the y-axis. The image of the given figure is a segment on the y-axis.

C sends (x,y) to (x,−y) so C reflects things in the x-axis.



Problem 3(a)

(b) The opposite of expanding radially is contracting radially so

$A^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$

The operator B isn't one-to-one since many points project to the same image on the y-axis. No inverse.

The opposite of reflecting in the x-axis is to reflect back again so $C^{-1} = C$.

(c) B projects onto the y-axis.

B^2 projects twice. But after a point is projected once, it's already on the y-axis so the second projection doesn't move it any more. So B^2 is the same mapping as B. Similarly $B^{100} = B$.

(d) C reflects in the x-axis. C^2 reflects twice which means that (x,y) is eventually sent back to (x,y) again. So $C^2 = I$.

(e) The range is the set of outputs.

With A, every point is an output; in particular (a,b) is the image of $(\frac{1}{2}a, \frac{1}{2}b)$.

The range of A is R^2.

With B, the outputs are all the points on the y-axis.
The range of B is the y-axis.
With C, every point is an output, corresponding to its reflection as input.
The range of C is R^2.

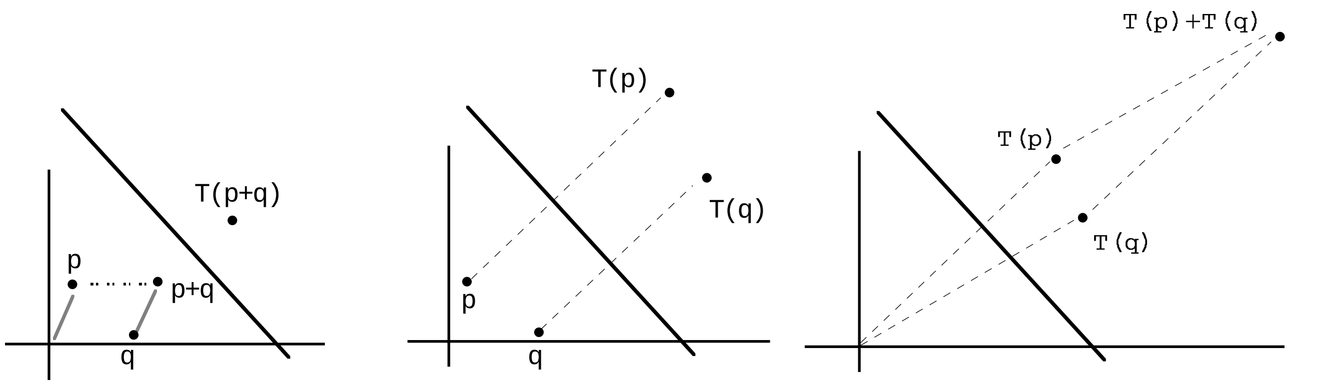
4. The vector u = (1,7) points along the line.
The line is a 1-dim subspace of R^2 with basis u.

I want the projection of point $\vec{x} = (x,y)$ onto the line:

$$\vec{x}_p = \frac{u \cdot x}{u \cdot u} u = \frac{x+7y}{50} u = \left(\frac{x+7y}{50}, \frac{7x+49y}{50}\right)$$

So M sends (x,y) to $\left(\frac{x+7y}{50}, \frac{7x+49y}{50}\right)$. So $M = \begin{bmatrix} 1/50 & 7/50 \\ 7/50 & 49/50 \end{bmatrix}$.

5. (a) The lefthand diagram shows p+q and T(p+q).
The middle diagram shows T(p) and T(q).
The righthand diagram shows T(p) + T(q).
(b) The two points T(p+q) in the lefthand diagram and T(p)+T(q) in the righthand diagram are not the same.
So T does not have the property T(u+v) = T(u)+T(v) for *all* u,v (my points p and q are a counterexample).
So T is not linear.



Problem 5

6. Let

$$A = \begin{bmatrix} \cos 60^\circ & -\sin 60^\circ \\ \sin 60^\circ & \cos 60^\circ \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then A rotates and B reflects so BA first rotates and then reflects. Answer is

$$BA = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}$$

7. No. The vectors that map to v are the solutions to the system of equations $M\vec{x} = \vec{v}$. The system can have no solutions, one solution or infinitely many. So either nothing maps to v or one vector maps to v or infinitely many map to v. But v can't have exactly 13 pre-images.

8. (a) $A = \begin{bmatrix} 1 & 5 & -2 \\ 4 & 3 & 7 \\ 2 & \pi & 8 \end{bmatrix}$

(b) Let

$$G = [u\ v\ w] = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix} \text{ and } H = [p\ q\ r] = \begin{bmatrix} 5 & 1 & 3 \\ 3 & 2 & 3 \\ 2 & 1 & 0 \end{bmatrix}.$$

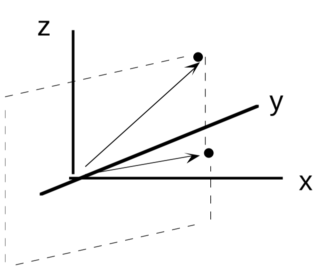
Then
G maps i,j,k to u,v,w respectively.
G^{-1} maps u,v,w to i,j,k respectively (inverse exists because u,v,w are ind).
H maps i,j,k to p,q,r respectively.
So HG^{-1} maps u,v,w to p,q,r. QED

9. (a) A maps (x,y,z) to (x,0,0) so A projects a point onto the x-axis.
Range of A is the x-axis, null space is the y,z plane (because those are the points that map to the origin), rank A = dim range A = 1.
B maps (x,y,z) to (x,y,0) so B projects a point onto the x,y plane.
Range of B is the x,y plane, null space is the z-axis, rank B = dim range B = 2.
C reflects a point in the x,y plane.
Range of C is R^3, null space contains only $\vec{0}$ and rank C = 3.

(b) Solution to = $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ is x=0; y,z free.

The set of points where x=0 and y and z can be anything is the y,z plane. So the null space is the y,z plane.

10. The cols of A are ind (no col is a combination of the preceding cols). So the range is a 3–dim subspace of R^3. So the range *is* R^3. The three cols of A are a basis but so are any 3 ind vectors in R^3, including i,j.,k.
The cols of B are dependent but the last two cols are ind. The range is a 2-dim subspace of R^3 with basis vectors u = (1,1,0), v = (1,1,1). The range is the plane in 3–space determined by points u,v and the origin (i.e., determined by arrows u and v attached to the origin).



Range of B
problem #10

11. (a) No because $Mu \neq \vec{0}$.

(b) *answer 1* Try to solve $M\vec{x} = \vec{u}$ to see if there is a pre-image $\vec{x}$. The system

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 3 \end{array}\right]$$

has no solution because the last two equations are z = 2 and 2z = 3 (no z can do that). So u is not in the range.
If you don't spot immediately that the system is inconsistent, you row op to echelon form

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 \end{array}\right]$$

The last row shows that the system is inconsistent.

answer 2 See if u is in the col space of M.
I can tell by inspection that it isn't; there is no way you can get
$a(1,0,0) + b(2,0,0) + c(1,1,2) = (1,2,3)$.

You would need $c = 2$ and also $c = 3/2$ to do it.

Or you can use row operations on $\begin{array}{cccc} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 3 \end{array}$ to get $\begin{array}{cccc} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 \end{array}$

In the echelon form, col 4 is not a combination of the other cols.
So back in the original, u is not a combination of the cols of M.

12. (a) Solve $M\vec{x} = \vec{0}$. The system row ops to $\left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

Solution is $x = -2z$, $y = -3z$, z free. The set of solutions is a 1-dim subspace of R^3. Any nonzero solution, say $(-2,-3,1)$, can serve as the one basis vector.

(b) *method 1* The range is the col space. To get a basis for the col space of M, do the row ops $R_1 \leftrightarrow R_2$, $R_3 = -2R_1 + R_3$, $R_4 = -R_1 + R_4$, $R_4 = -R_2 + R_4$ to get

$$\begin{matrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$$

Echelon cols 1 and 2 have pivots so cols 1 and 2 of the original matrix are a basis for the col space. So a basis for the range is u = (0,1,2,1), v = (1,0,0,1)

method 2 The range is the set of $\vec{b}$'s that make $M\vec{x} = \vec{b}$ consistent. So solve

$$\left[\begin{array}{ccc|c} 0 & 1 & 3 & b_1 \\ 1 & 0 & 2 & b_2 \\ 2 & 0 & 4 & b_3 \\ 1 & 1 & 5 & b_4 \end{array}\right]$$

Use the same row ops as method 1 to get

$$\left[\begin{array}{ccc|l} 1 & 0 & 2 & b_2 \\ 0 & 1 & 3 & b_1 \\ 0 & 0 & 0 & -2b_2 + b_3 \\ 0 & 0 & 0 & -b_1 - b_2 + b_4 \end{array}\right]$$

For the system to be consistent you need $-2b_2 + b_3 = 0$ and $-b_1 - b_2 + b_4 = 0$. The range is the set of points (b_1,b_2,b_3,b_4) such that $b_3 = 2b_2$ and $b_4 = b_1 + b_2$.

If you can't see a basis by inspection, write these equations as

$$\begin{aligned} b_1 &= t \\ b_2 &= s \\ b_3 &= 2s \\ b_4 &= s + t \end{aligned}$$

The range is the set of points of the form

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

and a basis is (1,0,0,1), (0,1,2,1).

Question I got different answers from methods 1 and 2. So isn't one of them wrong?
Answer No. They are both right. A subspace has lots of bases.

method 3 Take the cols of M, line them up as rows and row op into echelon form:

$$\begin{matrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{matrix}$$

The nonzero echelon rows are a basis for the original row space so they are a basis for the col space of M.

13. (a) M is invertible. Range is all of R^{10}, null space contains only $\vec{0}$, $|M| \neq 0$.
(b) Range is a 7-dim subspace of R^{10}, null space is a 3-dim subspace of R^{10}, M isn't invertible, $|M| = 0$.

14. (a) In the echelon form of M, 5 of the 9 cols have pivots. The range is a 5-dim

subspace of R^5 so the range *is* R^5.

Null space is a 4-dim subspace of R^9 because there are 4 free variables when you solve $M\vec{x} = \vec{0}$. Also because of the formula: dim of the null space is $n-r$ where $n = 9$, $r = 5$.

(b) In the echelon form of M, 3 of the 9 cols have pivots. Range is a 3-dim subspace of R^5. Null space is a 6-dim subspace of R^9.

15. (a) False. The range is a 4-dim subspace of R^6 which is not the same as R^4.

(b) True. The range is a 6-dim subspace of R^6, and the only 6-dim subspace of R^6 is R^6 itself.

16. (a) Not possible. First of all, $A\vec{0} = \vec{0}$ so $\vec{0}$ is always an output.
Secondly, the set of outputs (the range) is a subspace. And the set containing just b is not a subspace (not closed).

(b) This is like asking how many sols there are to the system $A\vec{x} = \vec{b}$.
Since we know that b *is* an output of A we know that there is at least one sol. The possibilities are either exactly one sol or infinitely many. So either there is just one input that produces b or there are infinitely many inputs that produce b.

(c) Now we know that the range of A is all of R^4. So A is invertible. So the system $A\vec{x} = \vec{b}$ has exactly one sol. So there is just one input that produces b.

17. (a) No. If M turns into M_1 after row ops, the systems of equations $M\vec{x} = \vec{0}$ and $M_1\vec{x} = \vec{0}$ have the same sols.

(b) Yes. The range is the col space, and row ops can change the col space (see row op rule (2) in Section 3.1).

18. (a) The set of vectors of the form (a,a,0,b) is a subspace with basis $u = (1,1,0,0)$, $v = (0,0,0,1)$.
The set of outputs of a matrix is the col space. So you want a matrix whose column space has basis u,v. One way to do it is to have cols u and v and then for the other two cols use any combinations of u and v so that u and v are a maximal ind set of cols. Or just make the last two cols $\vec{0}$.

There are lots of answers. One possibility is $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

(b) Not possible. The set of vectors of the form (a,a,2,b) is not a subspace (not closed under scalar multiplication) (not closed under addition) (doesn't contain $\vec{0}$). So it can't be the range of any matrix operator.

19. (a) $\text{Rank}(AB) \le 4$ and $\text{rank}(AB) \le 7$. So the best conclusion is $\text{rank}(AB) \le 4$.
(b) $\text{Rank}(AB) = 4$.
(c) Since you end up with a contradiction, what this proves is that A can't be invertible in the first place.
Here's another way to see it. If A were invertible with rank 4 then it would have to be 4×4. Since AB exists, B has to have 4 cols. But then it can't have rank 7. So A can't be invertible.

20 . The matrix I in question must be 3×3. And B must be 4×3.
$r(AB) = r(I) = 3$
$r(AB) \le r(B)$ (rule about rank of a product)
So $r(B) \ge 3$.
But rank B can't be > 3 since B only has 3 cols. So rank B = 3.

21. (a) The range of A is a 4-dim subspace of R^7

(b) The vectors $A\vec{x}_1, \ldots, A\vec{x}_6$ are in the range of A. So they are *six* vectors in a *four*-dim subspace of R^7. But 6 vectors in a 4-dim subspace must be dep. QED.

22. Let $A = \begin{bmatrix} 4 & 6 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$.

Then

A sends i,j to u,p respectively
B sends i,j to v,q respectively
A^{-1} sends u,p to i,j respectively

So BA^{-1} sends u,p to v,q. So

$$M = BA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \times \frac{1}{4} \begin{bmatrix} 1 & -6 \\ 0 & 4 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1/4 & 2/4 \\ 3/4 & 2/4 \end{bmatrix}$$

check $Mu = \begin{bmatrix} 1/4 & 2/4 \\ 3/4 & 2/4 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = v$

$Mp = \begin{bmatrix} 1/4 & 2/4 \\ 3/4 & 2/4 \end{bmatrix} \begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix} = q$ QED

SOLUTIONS Section 6.2

1. Let $P = \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix}$. Then $P^{-1} = \frac{1}{6}\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$.

New matrix for the trans is $P^{-1}MP = \begin{bmatrix} 5/2 & -1/2 \\ 9/2 & -1/2 \end{bmatrix}$.

Trans sends (x,y) to (x, 3x+y) and sends ((X,Y)) to $((\frac{5}{2}X - \frac{1}{2}Y,\ \frac{9}{2}X - \frac{1}{2}Y))$.

2. (a) Let $P = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}$. Then $P^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -2 \\ 0 & 2 \end{bmatrix}$ and

$M = P^{-1}AP$

$A = PMP^{-1} = \begin{bmatrix} 7 & -14 \\ 2 & -5 \end{bmatrix}$

(b) $\begin{bmatrix} 3 & 1 \\ 4 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$ so T(u+3v) = 6u + v.

(c) Convert u+3v to the i,j basis.
u+3v = (2,0) + 3(2,1) = (8,3).

$\begin{bmatrix} 7 & -14 \\ 2 & -5 \end{bmatrix}\begin{bmatrix} 8 \\ 3 \end{bmatrix} = \begin{bmatrix} 14 \\ 1 \end{bmatrix}$ so T(u+3v) = T(8i+3j) = 14i + j

Check that the answers to (b) and (c) agree:
6u+v = 6(2,0) + (2,1) = 14i + j. QED

3. (a) The coords of $\vec{u}$ w.r.t. basis $\vec{u},\vec{v},\vec{w}$ are 1,0,0 (because u = 1u + 0v + 0w).

$M\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$ so T(u) = 2u + 5v + 8w = −22i + 29j + 18k

Question What's wrong with finding $M\begin{bmatrix} 1 \\ 4 \\ 6 \end{bmatrix}$. It comes out to be $\begin{bmatrix} 38 \\ 71 \\ 50 \end{bmatrix}$ so why isn't the answer T(u) = (38,71,50).

Answer Two things wrong here.
Remember that M is the matrix of the transformation w.r.t. basis u,v,w. So when you find $M\begin{bmatrix} 1 \\ 4 \\ 6 \end{bmatrix}$ it means that you are finding T(u+4v+6w). But the problem was to find T(u) not T(u+4v+6w). And furthermore the 38,71,50 you get out means that the image of u+4v+6w is 38u+71v+50w, not (38,71,50).

(b) Let $P = [u\ v\ w] = \begin{bmatrix} 1 & -8 & 2 \\ 4 & 1 & 2 \\ 6 & -2 & 2 \end{bmatrix}$

I'll think of M as the new matrix of T. I want the old matrix (i.e., w.r.t. i,j,k).
new = P^{-1} old P so the old matrix is PMP^{-1}.

As a check, I used Mathematica to compute PMP^{-1} and then I found $PMP^{-1}\begin{bmatrix} 1 \\ 4 \\ 6 \end{bmatrix}$ and I got $\begin{bmatrix} -22 \\ 29 \\ 18 \end{bmatrix}$ as I should.

4. *step 1* One possibility is u = (1,3), v = (−3,1)
step 2 The reflection of ((X,Y)) is ((X,−Y)). You want a matrix B such that

$$B \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X \\ -Y \end{bmatrix}. \text{ So } B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

step 3 Let $P = \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$. Then $P^{-1} = \begin{bmatrix} 1/10 & 3/10 \\ -3/10 & 1/10 \end{bmatrix}$, $P^{-1}AP = B$ so

$$A = PBP^{-1} = \begin{bmatrix} -8/10 & 6/10 \\ 6/10 & 8/10 \end{bmatrix}$$

check: $A \begin{bmatrix} -5 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$ so the reflection of $(-5,5)$ in the line $y=3x$ is $(7,1)$. If you plot the points, they do look like reflections.

5.(a) $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

(b) Let $P = \begin{bmatrix} 1 & -5 \\ 4 & 2 \end{bmatrix}$. Then $B = P^{-1}AP = \frac{1}{22} \begin{bmatrix} -18 & -20 \\ -8 & 18 \end{bmatrix}$

(c) $B \begin{bmatrix} X \\ Y \end{bmatrix} = \frac{1}{22} \begin{bmatrix} -18 & -20 \\ -8 & 18 \end{bmatrix} = \begin{bmatrix} -\frac{18}{22} X - \frac{20}{22} Y \\ -\frac{8}{22} Y + \frac{18}{22} Y \end{bmatrix}$

so T sends $((X,Y))$ to $((-\frac{18}{22} X - \frac{20}{22} Y, -\frac{8}{22} X + \frac{18}{22} Y))$

6. (a) Points like (3,2) and (2,3) are reflections of one another in the line $y=x$
In general, T reflects points in the line $y=x$ (see the diagram).

(b) By inspection, the matrix of T w.r.t. basis i,j is $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

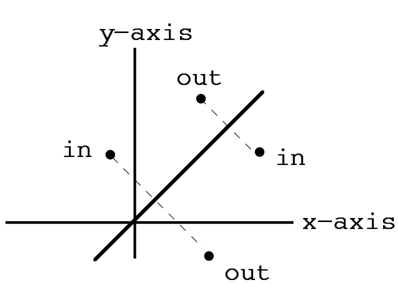
Let $P = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$.

The matrix for T w.r.t. basis u,v is

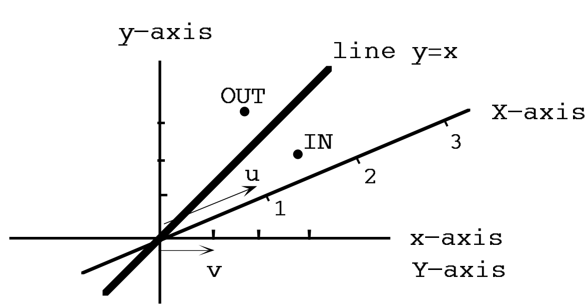
$$P^{-1}AP = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -3 & 2 \end{bmatrix}$$

(c) $\begin{bmatrix} 2 & 1 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$ so $T(2u-v) = 3u-4v$.

The diagram shows the X,Y coord system with basis u,v. The Y-axis is the same as the X-axis.
I plotted the input point ((2,−1)) and the output point ((3,−4)). They do look like reflections in the line $y=x$.
As a further check, $2-v = (3,2)$ and $3u-4v$ does out to be (2,3), as it should.



Problem 6(a)



Problem 6(c)

7. (a) Let $P_1 = [p\ q\ r] = \begin{bmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{bmatrix}$

Similarly, let $P_2 = [u\ v\ w]$.

Let M_{ijk} be the matrix of the trans w.r.t. basis i,j,k.

Then

$$M_{pqr} = P_1^{-1} M_{ijk} P_1$$

$$M_{uvw} = P_2^{-1} M_{ijk} P_2$$

So

$$M_{ijk} = P_1 M_{pqr} P_1^{-1}$$

$$M_{uvw} = P_2^{-1} P_1 M_{pqr} P_1^{-1} P_2$$

(b) I want to write M_{uvw} as $Q^{-1} M_{pqr} Q$. From part (a), I have

$$M_{uvw} = P_2^{-1} P_1 M_{pqr} P_1^{-1} P_2$$

But $P_2^{-1} P_1$ can be written as $(P_1^{-1} P_2)^{-1}$ by inverse rules so

$M_{uvw} = (P_1^{-1} P_2)^{-1} M_{pqr} P_1^{-1} P_2$ which is of the form $Q^{-1} M_{pqr} Q$ where $Q = P_1^{-1} P_2$.

8. Suppose A and B are similar. Then $B = P^{-1}AP$ for some P.
But multiplying by invertible matrices (namely P^{-1} and P) doesn't change rank.
So B has same rank as A.

9. There is a matrix Q such that $A = Q^{-1} BQ$ because A and B are similar. So

$A^T = (Q^{-1} BQ)^T$ take T on both sides

$= Q^T B^T (Q^{-1})^T$ T rule

$= Q^T B^T (Q^T)^{-1}$ inverse rule

There are two ways to continue now.

version 1 Solve for B to get

$$B^T = (Q^T)^{-1} A^T Q^T$$

This makes A^T and B^T similar because $B^T = R^{-1} A^T R$ where R is Q^T.

version 2 Rewrite as

$$A^T = \quad = \left[(Q^T)^{-1} \right]^{-1} B^T (Q^T)^{-1}$$

This makes A^T and B^T similar because $A^T = P^{-1} B^T P$ where $P = (Q^T)^{-1}$.

SOLUTIONS review problems for Chapters 5 and 6

1.(a) Start with
$$\left[\begin{array}{rrr|r} 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{array}\right]$$

Do row ops
row2 = −row1 + row2
row3 = −row1 + row3
row4 = −row1 + row4
row4 = −row2 + row4
row4 = row3 + row4

to get
$$\left[\begin{array}{rrr|r} 1 & -1 & -1 & 1 \\ 0 & 2 & 2 & -1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & -1 \end{array}\right]$$

The last line makes the system inconsistent.

(b) The vector $\vec{i} = (1,0,0,0)$ is not in the subspace spanned by $\vec{u_1} = (1,1,1,1)$, $\vec{u_2} = (-1,1,-1,1)$, $\vec{u_3} = (-1,1,1,-1)$.

(c) *method 1* (better)

The least squares solution is the projection of $\vec{i}$ onto the col space.

The cols $\vec{u_1}$, $\vec{u_2}$, $\vec{u_3}$ are orthogonal so

$$\vec{i}_{\text{proj}} = \frac{u_1 \cdot i}{u_1 \cdot u_1} u_1 + \frac{u_2 \cdot i}{u_2 \cdot u_2} u_2 + \frac{u_3 \cdot i}{u_3 \cdot u_3} u_3 + \frac{u_4 \cdot i}{u_4 \cdot u_4} u_4$$
$$= \frac{1}{4}\vec{u_1} - \frac{1}{4}\vec{u_2} - \frac{1}{4}\vec{u_3}$$

So the least squares sol to the inconsistent system is $a = 1/4$, $b = -1/4$, $c = -1/4$.

method 2

Let $A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$. The normal equations are

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

Least squares solution is $a = 1/4$, $b = -1/4$, $c = -1/4$.

(d) The original system asked you to find a,b,c so that $a\vec{u_1} + b\vec{u_2} + c\vec{u_3}$ is $\vec{i}$. It turned out to be impossible but using $a = 1/4$, $b = -1/4$, $c = -1/4$ will at least make $\|au_1 + bu_2 + cu_3 - i\|$ minimum (i.e., they make the vector $au_1+bu_2+cu_3$ as close to i as possible).

2. $A = Q^{-1}BQ$ for some Q.
Q,B and Q^{-1} are invertible and the product of invertibles is invertible. So A is invertible.

3. $|M| = 0$, rank M = 0, M maps everything to $\vec{0}$ so range contains only $\vec{0}$ and the null space is all of R^4.

4. There are many answers. Here are some possibilities.

(a) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

There is one solution if $\vec{b} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$ and no solutions if $\vec{b} = \begin{bmatrix} 6 \\ 7 \\ 8 \end{bmatrix}$.

(b) Let $A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \end{bmatrix}$

If $\vec{b} = \begin{bmatrix} a \\ b \end{bmatrix}$ then the solution is $x_1 = a - x_3$, $x_2 = b - 4x_3$; x_3 free.

(c) Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

If $\vec{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ then there are no solutions. If $\vec{b} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$ then there are infinitely many solutions, namely $x_1 = 2 - 2x_2$, $x_3 = 3$, x_2 free.

(d) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (any invertible matrix will work).

(e) Not possible. You can't avoid having a solution when $\vec{b} = \vec{0}$ because $Ax = \vec{0}$ always has at least the trivial solution.

5.(a) rank AB = 6 (b) rank AB $\le$ 6

6. (a) rank = number of echelon cols with pivots = 3

(b) The range is the col space. Basis vectors are the original cols 1,4,6 in M. Col space is 3-dim.

(c) Solve $M\vec{x} = \vec{0}$. Sol is

$$\begin{aligned} x_6 &= 0 \\ x_4 &= -6x_5 \\ x_1 &= -5x_5 - 3x_3 - 2x_2 \\ &x_2, x_3, x_5 \text{ free} \end{aligned}$$

Set $x_2 = 1$, $x_3 = 0$, $x_5 = 0$ to get $u = (-2,1,0,0,0,0)$.
Set $x_2 = 0$, $x_3 = 1$, $x_5 = 0$ to get $v = (-3,0,1,0,0,0)$.
Set $x_2 = 0$, $x_3 = 0$, $x_5 = 1$ to get $w = (-5,0,0,-6,1,0)$.
Basis for the null space is u,v,w. The null space is 3-dim.

(d) No. For one reason, $\vec{u}$ doesn't satisfy the system $M\vec{x} = \vec{0}$ because it doesn't satisfy the equivalent system $(\text{echM})\vec{x} = \vec{0}$. For another reason, part (c) gave the solution to $Mx = \vec{0}$ and u doesn't fit (among other things, its 6th component isn't 0).

(e) 0. Since rank M = 3 , the largest order of a nonzero subdet is 3. Subdets larger than 3×3 must be 0.

(f) Can't do. At least one 3×3 subdet is nonzero but you can't tell which one(s).

(g) Can't do it without knowing what row ops it took to get M into echelon form.

To solve $M\vec{x} = \vec{b}$, I want to do row ops not just to M but to $M|b$. Unless I know what the row ops do to b, I can't solve the system of equations.

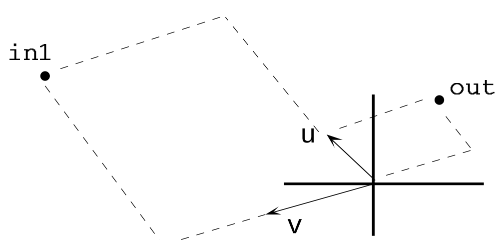
7. Let P be the basis changing matrix whose columns are the i,j,k coordinates of the new basis vectors. Then

$$\begin{aligned} B &= P^{-1}AP \\ |B| &= |P^{-1}AP| \\ &= |P^{-1}|\ |A|\ |P| && \text{det rule} \\ &= |A| && \text{because } |P^{-1}| = \frac{1}{|P|} \end{aligned}$$

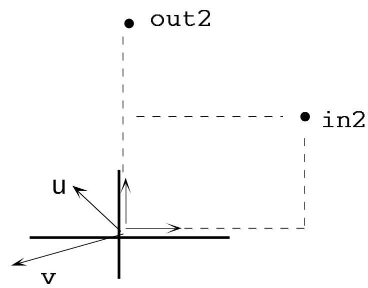
Note: $P^{-1}AP$ does not equal $P^{-1}PA$; you can't change the order of the factors in a matrix product. So in the matrix product $P^{-1}AP$, the P^{-1} and P do not cancel out. But $|P^{-1}||A||P|$ does equal $|P^{-1}||P||A|$ because determinants are numbers and you can change the order of the factors in a product of *numbers*. So in the product $|P^{-1}||A||P|$, $|P^{-1}|$ and $|P|$ do cancel out.

8. The rank of M is either 0 or 1 but not 2. So the range of M is either a 0-dim or a 1-dim subspace of R^2. So at best, the outputs of M can spread out along a line through the origin. But they can't spread out as much as the triangular region in the righthand diagram. The correct image of the lefthand region would have to lie along a line through the origin.

9. (a) The lefthand diagram shows the input in1 = 3u + 2v and the corresponding output out1 = u − v.
(b) The righthand diagram shows the input in2 = 3i + 2j and the corresponding output out2 = 3j.
I would have put all of this in one picture but it got too crowded.
Note: There is no connection between in1 and in2 or between out1 and out2. There was never intended to be any connection.
(c) Let the columns of P be the coordinates of u and v (w.r.t. i,j). Then
$A = P^{-1}BP$



Problem 9(a)



9(b)

10. Find a basis for the orthog complement of the space spanned by u,v,w.
Let A have rows u,v,w. Solve the system of equations $A\vec{x} = \vec{0}$
(Equivalently, solve the system $\vec{u}\cdot\vec{x} = 0$, $\vec{v}\cdot\vec{x} = 0$, $\vec{w}\cdot\vec{x} = 0$.)

$$\left[\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 0 \\ 2 & 6 & 9 & 5 & 0 \\ -1 & -3 & 3 & 0 & 0 \end{array}\right] \quad \text{row ops to} \quad \left[\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Sol is

$$\begin{aligned} x_3 &= -\tfrac{1}{3}x_4 \\ x_1 &= -x_4 - 3x_2 \end{aligned}$$

Set $x_2 = 1$, $x_4 = 3$ (avoids fractions) to get solution p = (−3,0,−1,3)

Set $x_2 = 1$, $x_4 = 0$ to get solution $q = (-3,1,0,0)$
A basis for the space of solutions is p,q.
So you can take p and q as the answer.

SOLUTIONS Section 7.1

1. $|z| = \sqrt{25 + 9} = \sqrt{34}$

$\overline{z} = 5 + 3i$

$zw = 2 - 42i$

$\frac{1}{z} = \frac{1}{5-3i} \cdot \frac{5+3i}{5+3i} = \frac{5+3i}{34} = \frac{5}{34} + \frac{3}{34}\,i$

$z^2 = 16 - 30i$

2. z must be real.

3. z must be pure imag (e.g., 3i, 6i, −2i).

4. $A^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, $A^3 = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}$, $A^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $A^5 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = A$

The products repeat with period 4.

By A^{243} you've gone through 60 cycles plus 3 rounds (because $243 = 60\cdot4 + 3$) so you land back at A^3. So $A^{243} = A^3 = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}$, $A^{244} = A^4$, $A^{245} = A$.

5. (a) $\sqrt{53}$ (b) 1 (c) 3 (d) 2 (e) 7 (f) 1 (g) $\sqrt{41}$

(h) $\sqrt{\cos^2\theta + \sin^2\theta} = 1$

6. (a) $zw = (2-i)(-3+4i) = -2 + 11i$, $\overline{zw} = -2 - 11i$

$\overline{z}\,\overline{w} = (2+i)(-3-4i) = -2 - 11i$

(b) $zw = -2 + 11i$, $|zw| = \sqrt{125} = 5\sqrt{5}$

$|z||w| = \sqrt{5}\,\sqrt{25} = 5\sqrt{5}$

7. You don't have to actually multiply $\pi + \sqrt{2}\,i$ and $\sqrt{17} + i$ together. By (5), the magnitude of the product is the product of the separate magnitudes. So

mag of the product = mag of z $\cdot$ mag of $\sqrt{17} + i$

$= \sqrt{\pi^2 + 2} \cdot \sqrt{18}$

SOLUTIONS Section 7.2

1. $A^* = \begin{bmatrix} 2 & 3+2i \\ i & 4-6i \end{bmatrix}$

2.(a) $6-2i$ (b) 7 (c) $6i$

3. $(A^*BC)^* = C^*B^*A^{**} = C^*B^*A$

4. $A^* = \begin{bmatrix} \bar{a} & \bar{d} & \bar{g} \\ \bar{b} & \bar{e} & \bar{h} \\ \bar{c} & \bar{f} & \bar{k} \end{bmatrix}$. To find $|A^*|$ I'll expand down col 1.

$$|A^*| = \bar{a}(\bar{e}\,\bar{k} - \bar{f}\,\bar{h}) - \bar{b}(\bar{d}\,\bar{k} - \bar{f}\,\bar{g}) + \bar{c}(\bar{d}\,\bar{h} - \bar{e}\,\bar{g})$$

$$= \overline{a(ek - fh) - b(dk - fg) + c(dh - eg)} \qquad \text{by conjugate rules}$$

The expression under the conjugate sign on the righthand side happens to be $|A|$ expanded across row 1. So

$$|A^*| = \overline{\text{det of } A} \qquad \text{QED}$$

5. (a) True. The diagonal entries are still real and matching entries $a+bi$ and $a-bi$ are now $-a-bi$ and $-a+bi$ which are still conjugates.
 (b) True. H^* *is* H so of course it's Herm.

6. *method 1* (good for 3×3's only)
 Let

$$A = \begin{bmatrix} a1 + a2\,i & a2 + b2\,i & a3 + b3\,i \\ a4 + b4\,i & a5 + b5\,i & a6 + b6\,i \\ a7 + b7\,i & a8 + b8\,i & a9 + b9\,i \end{bmatrix}$$

where $a1, \dots, a9, b1, \dots, b9$ are all real.
Then

$$A + A^* = \begin{bmatrix} 2a1 & a2+a4 + (b2-b4)i & \text{etc} \\ a2+a4 - (b2-b4)i & 2a5 & \text{ec} \\ \text{etc} & \text{etc} & 2a9 \end{bmatrix}$$

You can see that $A+A^*$ is Herm.
 method 2 (good for $n \times n$'s in general)
 I'll show that $(A + A^*)^* = A + A^*$.

$$\begin{aligned} (A + A^*)^* &= A^* + A^{**} && * \text{ rule} \\ &= A^* + A && * \text{ rule} \\ &= A + A^* \end{aligned}$$

7. Want to show that $(A^*HA)^* = A^*HA$.
 $(A^*HA)^* = A^*H^*A^{**} = A^*H^*A = A^*HA$ (by $*$ rules and the hypothesis that H is Herm)

8. Want to show that $(H^{-1})^* = H^{-1}$.

$$\begin{aligned} (H^{-1})^* &= (H^*)^{-1} && * \text{ rule} \\ &= H^{-1} && H \text{ is Herm so } H^* = H \end{aligned}$$

9. *method 1* $(A - A^*)^* = A^* - A^{**} = A^* - A = -(A - A^*)$ so $A - A^*$ is skew Herm.
 method 2 (good for 3×3's) Take the typical complex matrix A from method 1 of #6 and compute $A - A^*$ to see that it is skew Herm.

10. $(K^2)^* = (KK)^* = K^*K^* = (-K)(-K) = K^2$ so K^2 is Herm.
 $(K^3)^* = (KKK)^* = K^*K^*K^* = (-K)(-K)(-K) = -K^3$ so K^3 is skew Herm.

SOLUTIONS Section 7.3

1. No since $u \cdot v = (-i)(-1) + (1)(i) = 2i$, not 0 (remember to conjugate the comonents of the first vector before you multiply and add).

2. $iu = (1+2i,\ 3i,\ 4)$

$u \cdot v = (2+i)(3-i) - 3i + 4i(1-2i) = 15 + 2i$ **warning** *Don't write* $u \cdot v = \bar{u} \cdot v$

$v \cdot u = \overline{u \cdot v} = 15 - 2i$

$\|u\| = \sqrt{4 + 1 + 9 + 16} = \sqrt{30}$

$\|v\| = \sqrt{9 + 1 + 1 + 1 + 4} = 4$

$|u \cdot v|$ (meaning the mag of the complex number $u \cdot v$) $= \sqrt{225 + 4} = \sqrt{229}$

$v_{unit} = \left(\frac{3-i}{4},\ -\frac{i}{4},\ \frac{1-2i}{4} \right)$

3. (a) $\sqrt{1 + 1} = \sqrt{2}$ (b) $\sqrt{4 + 9 + 1} = \sqrt{14}$ (c) $\sqrt{9 + 1 + 4 + 16} = \sqrt{30}$

4.(a)
$$
\begin{aligned}
(u+v)\cdot(u-v) &= u\cdot u - v\cdot v + v\cdot u - u\cdot v \\
&= \|u\|^2 - \|v\|^2 + v\cdot u - u\cdot v \\
&= 9 - 49 + 6 + 2i - (6-2i) \\
&= -40 + 4i
\end{aligned}
$$

(b)
$$
\begin{aligned}
\|u + iv\|^2 &= (u + iv)\cdot(u + iv) \\
&= u\cdot u + iv\cdot u + u\cdot iv + iv\cdot iv \\
&= u\cdot u + \bar{i}(v\cdot u) + i(v\cdot u) + \bar{i}\,i(v\cdot v) \\
&= \|u\|^2 - i(v\cdot u) + i(u\cdot v) + \|v\|^2 \\
&= 9 - i(6 + 2i) + i(6 - 2i) + 49 \\
&= 62
\end{aligned}
$$

$\|u + iv\| = \sqrt{62}$

> **warning**
> $\|u + iv\|^2$ is $(u+iv)\cdot(u+iv)$.
> It is *not* $(u-iv)\cdot(u+iv)$.

(c) $(2-3i)u\cdot iu + v\cdot iu = \overline{2-3i}\ i(u\cdot u) + i(v\cdot u)$

$= (2+3i)i\|u\|^2 + i(6 + 2i) = -29 + 24i$

(d) $\|6iu\| = |6i|\,\|u\|$ (where $|6i|$ means the mag of the number $6i$)

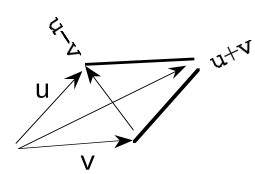
$= 6 \times 3 = 18$

(e) $\|(2-3i)u\| = |2-3i|\,\|u\| = 3\sqrt{13}$

5. (a) True

(b) False. It multiplies the norm by the mag of 3i which is 3.

6. (a) True because the diagonals of a rhombus are perp.



Problem 5(a)

(b) If $\|u\| = \|v\|$ then
$$
\begin{aligned}
(u+v)\cdot(u-v) &= u\cdot u - v\cdot v \\
&= \|u\|^2 - \|v\|^2 \\
&= 0
\end{aligned}
$$
so $u+v$ and $u-v$ are orthog.

(c) The step $(u+v)\cdot(u-v) = u\cdot u - v\cdot v$ breaks down in C^n. What you have instead is $(u+v)\cdot(u-v) = u\cdot u + v\cdot u - u\cdot v + v\cdot v$.

In R^n, $u\cdot v$ and $v\cdot u$ are equal so those terms cancel out. But they aren't necessarily equal in C^n.

For a counterexample, choose u and v so that $u\cdot v$ is non-real so that it doesn't equal $v\cdot u$. For instance let $u = (i,0)$ and $v = (1,0)$. Then $\|u\| = \|v\|$ (both are 1) but

$(u+v)\cdot(u-v) = (1+i,\ 0)\cdot(-1+i,\ 0) = (1-i)(-1+i) = 2i$, not 0.

7. $\|x + iy\|^2 = (x + iy)\cdot(x + iy)$
$= x\cdot x + (iy)\cdot(iy) + (iy)\cdot x + x\cdot(iy)$
$= x\cdot x + (-i)(i)(y\cdot y) + (-i)(y\cdot x) + i(x\cdot y)$
$= \|x\|^2 + \|y\|^2 + i(x\cdot y - y\cdot x)$

8. Let $x\cdot y = a + bi$. Then $y\cdot x = a - bi$ and

$\frac{1}{2}i(\vec{y}\cdot\vec{x}) - \frac{1}{2}i(\vec{x}\cdot\vec{y}) = \frac{1}{2}i(a-bi) - \frac{1}{2}i(a+bi) = b$ which is the imag part of $x\cdot y$

9. (a) Let $x = au + bv$. Then

$(2+3i,\ 6-7i) = a(i,0) + b(0,i)$

$2+3i = ia,\ a = \frac{2+3i}{i} = -2i+3$ (note that $1/i = -i$)

$6-7i = ib,\ b = \frac{6-7i}{i} = -6i - 7$

(b) $P = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}$, $P^{-1} = \begin{bmatrix} -i & 0 \\ 0 & -i \end{bmatrix}$, $P^{-1}\begin{bmatrix} 2+3i \\ 6-7i \end{bmatrix} = \begin{bmatrix} 3-2i \\ -7-6i \end{bmatrix}$

so the new coords of $\vec{x}$ are $3-2i$ and $-7-6i$

(c) u and v are orthog and unit vectors so

$x = (u\cdot x)u + (v\cdot x)v = (3-2i)u + (-7-6i)v$

warning It's OK to use the formula $x = \frac{u\cdot x}{u\cdot u}u + \frac{v\cdot x}{v\cdot v}v$ but make sure you get $u\cdot u = \bar{i}\, i = 1$. It is *wrong* to say $u\cdot u = i^2 = -1$.

10. It computes $v\cdot u$ or equivalently, $-u\cdot v$.

SOLUTIONS Section 7.4

1. Let U_1 and U_2 be unitary. Want to show that $(U_1U_2)(U_1U_2)^* = I$.

$$
\begin{aligned}
(U_1U_2)(U_1U_2)^* &= U_1U_2\, U_2^*\, U_1^* && \text{* rule} \\
&= U_1\, U_1^* && U_2U_2^* = I \text{ since } U_2 \text{ is unitary} \\
&= I && \text{since } U_1 \text{ is unitary}
\end{aligned}
$$

2. The cols of U are orthonormal. So the rows of U^T are orthonormal so U^T is also unitary.

3. (a) Not unitary since $||M|| = 2$ not 1.
(b) $||M|| = 1$ which is inconclusive.
(c) Not unitary since $||M| = \sqrt{5}$ not 1.

(d) $||M|| = 1$, inconclusive.

(d) $||M|| = 1$, inconclusive.

(e) $||M|| = 1$, inconclusive.

4. Not unitary since the cols are no longer unit vectors (they have norm 3 now).

5.
$$
\begin{aligned}
(U^{-1})\ (U^{-1})^* &= U^{-1}\ (U^*)^{-1} && \text{* rule} \\
&= U^{-1}\ (U^{-1})^{-1} && \text{since U is unitary} \\
&= U^{-1}\ U \\
&= I
\end{aligned}
$$

6. If $U_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $U_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ then $U_1 + U_2 = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$.

U_1 and U_2 are unitary but $U_1 + U_2$ is not.

SOLUTIONS review problems for Chapter 7

1. $u \cdot v = (-2i)(3+4i) + (-6)(5+6i) = -22 - 42i$

$v \cdot u = \overline{u \cdot v} = -22 + 42i$

$\|u\| = \sqrt{4 + 36} = \sqrt{40}$

$\|v\| = \sqrt{9 + 16 + 25 + 36} = \sqrt{86}$

2. Let the scalar $w \cdot v$ be called k temporarily. Then

$$
\begin{aligned}
(v - kw) \cdot w &= v \cdot w - (kw) \cdot w \\
&= v \cdot w - \overline{k}(w \cdot w) \\
&= v \cdot w - \overline{k}\ \|w\|^2 \\
&= v \cdot w - 9\overline{k} \\
&= v \cdot w - 9\ \overline{w \cdot v} \qquad \text{plug the k back in} \\
&= v \cdot w - 9\ (v \cdot w) \\
&= -8(v \cdot w)
\end{aligned}
$$

3. $H = H^*$ so $|H| = |H^*| = \overline{|H|}$.

So $|H|$ equals its conjugate. So $|H|$ is real.

4. Want to show that $(U^{-1}HU)^* = U^{-1}HU$.

$$
\begin{aligned}
(U^{-1}HU)^* &= U^* H^* (U^{-1})^* \qquad && * \text{ rule} \\
&= U^{-1} H (U^*)^* && \text{since U is unitary and H is Herm} \\
&= U^{-1}HU
\end{aligned}
$$

5. Let

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{and } B = \begin{bmatrix} 1 & 4 \\ 4 & 7 \end{bmatrix}$$

Then A and B are Herm but $AB = \begin{bmatrix} \cdot & 34 \\ 14 & \cdot \end{bmatrix}$ is not Herm.

6. (a) *method 1* Each col has norm 1. And the dot product of any col with any other col is 0. So the cols are orthonormal which makes A unitary.

method 2 Do the same thing as method 1 but with rows.

method 3

$$A^*A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & i \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & -i \end{bmatrix} = I$$

So A is unitary.

(b) A^{-1} is A^*, which I found above.

7. (a) Not possible. First col isn't a unit vector.

(b) Let $x = -3i$, y any real number

(c) Let $x = 3i$, y any pure imaginary.

8. $\|u\| = \sqrt{1 + 1 + 4 + 9 + 4}$ so $u_{\text{unit}} = \left(\frac{i}{\sqrt{19}}, \frac{i}{\sqrt{19}}, \frac{2}{\sqrt{19}}, \frac{2+3i}{\sqrt{19}} \right)$.

9. The diagonal entries are real by property (5) of dot products.
The matching entries like $u \cdot w$ and $w \cdot u$ are conjugates by property (1) of dot products.
So A is Hermitian.

SOLUTIONS Section 8.1

1. $Au = \begin{bmatrix} -6 \\ 6 \\ 0 \end{bmatrix} = -2u$ so u is an eigenvector corr to $\lambda = -2$.

$Av = \begin{bmatrix} 4 \\ 2 \\ \text{who cares} \end{bmatrix}$.

The output is not a multiple of the input so v is not an eigenvector.

2. (1) u is an eigenvector of M corr to $\lambda = 6$.
(2) Mu = 0u so u is an eigenvector of M corr to $\lambda=0$ (eigen*values* can be 0; it's eigen*vectors* that can't be $\vec{0}$).
(3) u is an eigenvector of AB corr to $\lambda = 3$.
(4) If $Bu \neq \vec{0}$ then Bu is an eigenvector of A corresponding to $\lambda = 3$.
(5) nothing
(6) u is an eigenvector of 2M corresponding to $\lambda = 1$. Also, you can write the equation as $Mu = \frac{1}{2}u$ so u is an eigenvector of M corr to $\lambda = \frac{1}{2}$.
(7) $M(u + v) = 3(u + v)$ so if $u+v \neq \vec{0}$ then u+v is an eigenvector corr to $\lambda = 3$.
(8) nothing
(This is a rule of algebra, but it doesn't say anything about eigenvectors.)
(9) u is an eigenvector of A+B corresponding to $\lambda = 2$.

3. Given $Mu = \lambda u$. Then $3Mu = 3\lambda u$ so $(3M)u = (3\lambda)u$.
This makes u an eigenvector of 3M with corresponding eigenvalue 3λ.

4. Iu = u = 1u so every nonzero vector is an eigenvector with corr eigenvalue 1.
There is one eigenspace, R^n itself.

5.(a) If u is in the null space then $Mu = \vec{0} = 0u$.
So every nonzero vector in the null space is an eigenvector corresponding to eigenvalue $\lambda = 0$.
(b) A *non*invertible matrix (see the invertible rule---for an invertible matrix, the null space contains only $\vec{0}$).

6. If u is an eigenvector of M corresponding to eigenvalue λ, so is any nonzero multiple of u.
u_{unit} is a nonzero multiple of u, namely $u_{unit} = \frac{1}{\|u\|} u$.

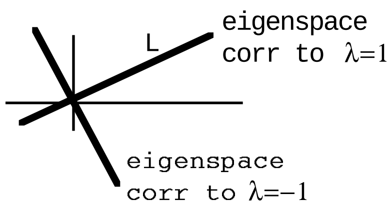
So u_{unit} is also an eigenvector corresponding to λ.

7. A sends points on the x-axis to themselves; i.e., if u is on the x-axis then Au = 1u. So the x-axis is an eigenspace corresponding to eigenvalue $\lambda = 1$.

And A sends points on the y-axis to $\vec{0}$, i.e., if u is on the y-axis then Au = 0u.
So the y-axis is an eigenspace corresponding to eigenvalue $\lambda = 0$.

B maps each point on L to itself. So line L is an eigenspace corresponding to $\lambda = 1$.
If u is on the line through the origin perp to L then B sends u to $-u$. So the line through the origin perp to L is an eigenspace corr to $\lambda = -1$ (see the diagram)

C sends every point u to 2u; i.e., for every vector u, Cu = 2u. So all of R^2 is an eigenspace with corr eigenvalue $\lambda=2$.



Problem 7B

8. (a) Yes. In fact each eigenvalue *must* have more than one eigenvector. It must have an entire eigenspace of eigenvectors. At the very least, if λ corresponds to eigenvector u then λ also has corresponding eigenvectors 2u, πu, $\frac{1}{2}$u etc.

(b) No. Suppose u is an eigenvector of M corresponding to eigenvalue $\lambda=2$ and also corresponding eigenvalue $\lambda=5$. Then $u \neq \vec{0}$ (eigenvectors are nonzero) and Mu = 2u and Mu = 5u. But then 2u = 5u which is impossible for a nonzero vector.

9. We know that $Au = \lambda u$ where $u \neq \vec{0}$. And $A = PBP^{-1}$ so

$PBP^{-1}u = \lambda u$

$BP^{-1}u = P^{-1}(\lambda u)$ left multiply by P^{-1}

$B(P^{-1}u) = \lambda(P^{-1}u)$ matrix algebra

That makes λ an eigenvalue of B corresponding to eigenvector $P^{-1}u$ provided that $P^{-1}u \neq \vec{0}$. And here's the proof that $P^{-1}u$ *is* nonzero.

The invertible rule (Section 6.1, item (7)) says that for an invertible matrix M if $x \neq \vec{0}$ then $Mx \neq \vec{0}$.

Apply this to P^{-1} and u: P^{-1} is invertible and $u \neq \vec{0}$ so $P^{-1}u \neq \vec{0}$.

10. Well, any two subspaces always overlap to some extent because they all must contain $\vec{0}$. But aside from having $\vec{0}$ in common, two eigenspaces can't have anything else in common. In other words, u can't be an eigenvector corresponding to say $\lambda=2$ and also corresponding to $\lambda=5$ as was shown in #8(b).

11. Since 0 is an eigenvalue, there is a *nonzero* vector $\vec{u}$ such that $M\vec{u} = 0\vec{u}$, i.e., such that $M\vec{u} = \vec{0}$. But the invertible rule (item (7), Section 6.1) says that this can't happen for an invertible matrix. So M *not* invertible.

SOLUTIONS Section 8.2

1. I'll call each matrix M.

(a) $|M-\lambda I| = \begin{vmatrix} -\lambda & 3 \\ 2 & -1-\lambda \end{vmatrix} = -\lambda(-1-\lambda) - 6) = \lambda^2 + \lambda - 6.$

Roots are $\lambda = -3,2$. Each is 1-fold so the matrix will have a complete set of eigenvectors.

To find an eigenvector for $\lambda = -3$, solve $(M - 3I)\vec{x} = \vec{0}$, $\left[\begin{array}{cc|c} 3 & 3 & 0 \\ 2 & 2 & 0 \end{array}\right]$

Sol is $x = -y$, y free. Set $y = 1$ to get eigenvector $(-1,1)$.

To find an eigenvector for $\lambda = 2$, solve $\left[\begin{array}{cc|c} -2 & 3 & 0 \\ 2 & -3 & 0 \end{array}\right]$

Sol is $x = \frac{3}{2}y$, y free. Set $y = 2$ to get eigenvector $(3,2)$.

(b) $\begin{vmatrix} i-\lambda & 1 \\ 1 & -i-\lambda \end{vmatrix} = \lambda^2$. The only eigenvalue is $\lambda = 0$ (2-fold).

To find eigenvectors, solve $\left[\begin{array}{cc|c} i & 1 & 0 \\ 1 & -i & 0 \end{array}\right]$

Both equations say $x = iy$. Eigenspace is 1-dim.
The matrix does not have a complete set of eigenvectors.
Set $y = 1$ to get eigenvector $(i, 1)$

(c) $|M-\lambda I| = \begin{vmatrix} 2-\lambda & 2 & -6 \\ 2 & -1-\lambda & -3 \\ -2 & -1 & 1-\lambda \end{vmatrix}$. Roots turn out to be $\lambda = -2,-2,6$.

To find eigenvectors for $\lambda = -2$, solve $\left[\begin{array}{ccc|c} 4 & 2 & -6 & 0 \\ 2 & 1 & -3 & 0 \\ -2 & -1 & 3 & 0 \end{array}\right]$

Each equation says $y = 3z - 2x$. There are 2 free variables so you can get two ind eigenvectors. The matrix is going to have a complete set of eigenvectors.
Set $x = 0$, $z = 1$, to get eigenvector $(0,3,1)$.
Set $x = 1$, $z = 0$, to get eigenvector $(1,-2,0)$.

To find an eigenvector for $\lambda = 6$ solve $\left[\begin{array}{ccc|c} -4 & 2 & -6 & 0 \\ 2 & -7 & -3 & 0 \\ -2 & -1 & -5 & 0 \end{array}\right]$

Row op to $\left[\begin{array}{ccc|c} -2 & 1 & -3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$

Solution is $y = -z$, $x = -2z$, z free. Set $z = 1$, to get eigenvector $(-2,-1,1)$.

(d) $|M-\lambda I| = -\lambda(1-\lambda)(-\lambda) - (1-\lambda)$ (I expanded down col 1)
Don't multiply out. It's smarter to pull out the common factor $1-\lambda$.

$$|M-\lambda I| = = (1-\lambda)(\lambda^2 - 1) = (1-\lambda)(\lambda-1)(\lambda+1)$$

Roots are $\lambda=1,1,-1$

To get eigenvectors corr to $\lambda=1$ solve $\left[\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right]$

Sol is y,z free, $x = z$. There are 2 ind eigenvectors, say $(1,0,1)$ and $(1,1,1)$.

To get eigenvectors corr to $\lambda=-1$ solve $\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$

Sol is $y=0$, z free, $x=z$. An eigenvector is $(1,0,-1)$.
The matrix has a complete set of eigenvectors.

(e) M is diagonal. By inspection, $M\mathbf{i} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} = 4\mathbf{i}$, $M\mathbf{j} = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix} = -5\mathbf{j}$, $M\mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ \pi \end{bmatrix} = \pi\mathbf{k}$

So 4 is an eigenvalue corr to eigenvector i, -5 is an eigenvalue with corr eigenvector j and π is an eigenvalue with corr eigenvector k.
The matrix has a complete set of eigenvectors.

2. (a) $|M-\lambda I| = \begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{vmatrix}$

$$= \cos^2\theta - 2\lambda\cos\theta + \lambda^2 + \sin^2\theta$$

$$= \lambda^2 - 2\lambda\cos\theta + 1$$

Solution to $\lambda^2 - 2\lambda\cos\theta + 1 = 0$ is

$$\lambda = \frac{2\cos\theta \pm \sqrt{4\cos^2\theta - 4}}{2}$$

$$= \frac{2\cos\theta \pm 2\sqrt{\cos^2\theta - 1}}{2}$$

$$= \frac{2\cos\theta \pm 2\sqrt{-\sin^2\theta}}{2}$$

$$= \cos\theta \pm i\sin\theta$$

(b) I'll get the eigenvector corresponding to $\lambda = \cos\theta + i\sin\theta$. Solve

$$\left[\begin{array}{cc|c} -i\sin\theta & -\sin\theta & 0 \\ \sin\theta & -i\sin\theta & 0 \end{array}\right]$$

Both equations simplify to $x = iy$. So an eigenvector is say $(i,1)$.

For the other eigenvalue, an eigenvector turned out to be $(-i,1)$.

3. Any nonzero multiple of u is also an eigenvector corresponding to $\lambda=0$.
And nonzero combination of v and w is also an eigenvector corresponding to $\lambda=9$.

4. Multiplicity is 2,3,4,5 or 6. Can't be 1 because a 1-fold λ has only one ind eigenvector.

5. Not possible. Every square matrix has eigenvalues and eigenvectors. A 5×5 matrix has 5 eigenvalues, counting multiplicity. And every eigenvalue has corresponding eigenvectors, a whole eigenspace full of them. The "worst" that could happen is that M has one 5−fold eigenvalue with only a 1−dim corresponding eigenspace so that even though there are infinitely many eigenvectors you can find only one *independent* eigenvector for poor M.

6. Characteristic poly of A has degree 8 so A is 8×8.
Eigenvalues are $\lambda=3$ (5-fold) with 5 or fewer corresponding ind eigenvectors
$\lambda=2$ (2-fold) with 1 or 2 ind eigenvectors
$\lambda=4$ (1-fold) with 1 ind eigenvector
Can't tell if A will have a complete set of eigenvectors.
Characteristic poly of B has degree 3 so B is 3×3.
Eigenvalues are $\lambda=-1,2,-3$ (each 1-fold). Each eigenvalue will have one ind eigenvector. B has a complete set of eigenvectors.

7. (a) $(-1,i)$ is a multiple of $(i,1)$ (the multiple is i) so $(-1,i)$ is also an eigenvector. Your answer is just as good as the book's answer.
(b) Yes, you are right. The key idea is that the set of eigenvectors corresponding to λ (together with the zero vector) is a *subspace* with basis u and v. Whether or not p is an eigenvector depends on whether or not p is *in* the subspace which in turn depends on whether or not p is a combination of u and v.
It turns out that $p = -u + 2v$. (One way to discover this is to solve the system of equations $au + bv = p$. Or you can line up u,v,p as cols and row op into echelon form.)

8. (a) 2 is an eigenvalue of A.
(b) 2 is not an eigenvalue of A.

9. (a) *by inspection* $Iu = u$ for every col vector u. So $\lambda=1$ is the only eigenvalue. And all the nonzero vectors in R^3 are corresponding eigenvectors. The eigenspace corresponding to $\lambda=1$ is 3-dim.

overkill $I - \lambda I = \begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{bmatrix}$, $|I-\lambda I| = (1-\lambda)^3$, $\lambda=1$ is a 3-fold eigenvalue.

The corresponding eigenvectors are the sols of $(I-1I)\vec{x} = \vec{0}$, But $I-1I$ is the zero matrix and all x's are solutions of the system. So all nonzero vectors are eigenvectors.

(b) *by inspection* If A is the zero matrix then $Au = \vec{0} = 0u$ for all u. So 0 is an eigenvalue and all nonzero vectors are corresponding eigenvectors. The eigenspace is 3-dim.

overkill $0-\lambda I = \begin{bmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{bmatrix}$, $|0-\lambda I| = (-\lambda)^3$, $\lambda=0$ is a 3-fold eigenvector.

To find eigenvectors, solve $(0-0I)\vec{x} = \vec{0}$. But $0-0I$ is the zero matrix and all x's are solutions. So all nonzero vectors are eigenvectors.

10. (a) $P^{-1}AP - \lambda I = P^{-1}(A - \lambda I)P$.

(b) Let A and B be similar. Then $B = P^{-1}AP$ for some invertible P.
I want to show that $|B-\lambda I| = |A-\lambda I|$.

$$\begin{aligned} |B - \lambda I| &= |P^{-1}AP - \lambda I| \\ &= |P^{-1}(A - \lambda I)P| && \text{by part (a)} \\ &= |P^{-1}||A - \lambda I|\,|P| && \text{by determinant rules} \\ &= |A - \lambda I| && |P^{-1}| \text{ and } |P| \text{ cancel since } |P^{-1}| = 1/|P| \end{aligned}$$

(c) The eigenvalues are the roots of the characteristic poly. A and B have the same characteristic poly so they must have the same eigenvalues.

11. (a) The eigenvectors are the solutions of $(A - 9I)\vec{x} = \vec{0}$.

A−9I has 5 cols and rank 2 so the set of solutions to $(A-9I)\vec{x} = \vec{0}$ is a 3-parameter family. So there are 3 ind eigenvectors.

(b) Infinitely many. The set of eigenvectors is a 3-dim subspace of R^5.

12. (a) When the matrix is non-invertible.

(b) The eigenvectors corresponding to $\lambda=0$ are the (non-zero) vectors satisfying the equation $M\vec{x} = 0\vec{x} = \vec{0}$. The eigenspace is the null space of M. Its dimension is n−r.

13. $Mu = \lambda u$ so $u = M^{-1}\lambda u = \lambda M^{-1} u$.

You know that $\lambda \neq 0$ since M is invertible so it's safe to divide by λ to get

$$M^{-1} u = \frac{1}{\lambda} u.$$

This makes u an eigenvector of M^{-1} with corresponding eigenvalue $1/\lambda$ provided $u \neq \vec{0}$. But u is an eigenvector of M. So it can't be $\vec{0}$. QED

14. Must have $|AB - 5I| = 0$.

15. (a) product of λ's (counting multiplicity) = det A = 4π

(b) sum of λ's (counting multiplicity) = trace A = $\sqrt{2} + 1$

16. (a) False. Let u = (1,0,0,0,0), v = (0,1,0,0,0), w = (0,0,1,0,0), p = (2,0,0,0,0), q = (0,3,0,0,0). Then u,v,w are ind; and p,q are ind. But u,v,w,p,q are dep.

(b) True. That's what it says in the paragraph "matrices with a complete set of eigenvectors" although I didn't prove it.

The moral is that merging independent bunches of vectors *might* result in a big *dependent* bunch. But merging ind bunches of eigenvectors corresponding to different eigenvalues results in a big *independent* bunch.

17. She can't be right. If you have a correct eigenvalue, there will *always* be corresponding eigenvectors (a whole eigenspace full of them). In other words, if λ really is an eigenvalue then the equation $(M - \lambda I)\vec{x} = \vec{0}$ will *always* have nonzero solutions. That's the whole point of the method for finding λ's in the first place.

So she must have made a mistake when she solved $(M - \lambda I)\vec{x} = \vec{0}$

SOLUTIONS Section 8.3

1. Eigenvectors of a Herm matrix corresponding to different eigenvalues are orthog. So $u \cdot v = 0$, $2 + 3y = 0$, $y = -2/3$.

2. Can't tell.
u and v are not orthogonal so if they correspond to *different* eigenvalues then M can't be Hermitian. But if they correspond to the *same* eigenvalue, M could be Hermitian.

3. (a) Can't tell.
   (b) Can't be Hermitian because the eigenvalues of a Hermitian must be real.

SOLUTIONS Section 8.4

1. (a) M is symmetric so it will be diagonalizable with an orthogonal P.

$$\begin{vmatrix} -\lambda & 2 \\ 2 & -\lambda \end{vmatrix} = \lambda^2 - 4. \quad \text{Roots are } \lambda = \pm 2.$$

To find an eigenvector for $\lambda = 2$, solve $\left[\begin{array}{cc|c} -2 & 2 & 0 \\ 2 & -2 & 0 \end{array}\right]$

Sol is x = y, y free. Set y = 1 to get eigenvector u = (1,1).

To find an eigenvector for $\lambda = -2$, solve $\left[\begin{array}{cc|c} 2 & 2 & 0 \\ 2 & 2 & 0 \end{array}\right]$

Sol is x = −y, y free. Set y = 1 to get eigenvector v = (−1,1).
One possible diagonalization is

$$P^{-1}MP = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \text{ where } P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

To diagonalize M with an *orthogonal* matrix, normalize u and v (they are already orthog). Then

$$Q^{-1}MQ = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \text{ where } Q = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

(b) $\begin{vmatrix} 2-\lambda & 1 \\ 0 & 2-\lambda \end{vmatrix} = (2-\lambda)^2$. So $\lambda=2$ is a 2-fold eigenvalue.

To find corresponding eigenvectors solve $\left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$

Sol is y = 0, x free.
Only one free variable so there is only one eigenvector plus multiples.
So M is not diagonalizable.

(c) To get eigenvectors for $\lambda = 2$, solve $\left[\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 \end{array}\right]$

Sol is x = z, y = z, z free. An eigenvector is u = (1,1,1).

To get eigenvectors corresponding to $\lambda = 1$, solve $\left[\begin{array}{ccc|c} -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \end{array}\right]$

Sol is x = y + z; y,z free.
There are two free variables so there are two ind eigenvectors so M is diagonalizable.
Set y = 1, z = 0 to get eigenvector v = (1,1,0).
Set y = 0, z = 1 to get eigenvector w = (1,0,1).
A diagonalization is

$$P^{-1}MP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ where } P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Don't see how to do it with a unitary matrix because the eigenvectors corr to $\lambda=1$ are not orthog to the eigenvector corr to $\lambda=2$.

2. (a) A is diagonalizable since each 2-fold λ produced 2 ind eigenvectors.
One diagonalization is

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \text{ where } P = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 2 & 1 \end{bmatrix}$$

Can be done with an orthogonal P since w and x are each orthogonal to u and v.
To get a set of orthogonal eigenvectors, use the Gram Schmidt process twice.

First use it on u,v. Let

$$u_1 = u = (1,0,0,0)$$

$$u_2 = v - \frac{u_1 \cdot v}{u_1 \cdot u_1} u_1 = (0,3,1,0)$$

And use it on w,x. Let

$$v_1 = w = (0,0,0,2)$$

$$v_2 = x - \frac{v_1 \cdot x}{v_1 \cdot v_1} v_1 = (0,1,-3,0).$$

Normalize u_1, u_2, v_1, v_2 and use them as cols of P. Then

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \quad \text{where } P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3/\sqrt{10} & 0 & 1/\sqrt{10} \\ 0 & 1/\sqrt{10} & 0 & -3/\sqrt{10} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(b) A can be diagonalized since each 2-fold eigenvalue has 2 ind eigenvectors. Don't see how to do it with a unitary P since the λ_2 eigenvectors are not both orthog to the λ_1 eigenvectors. One diagonalization is

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \text{where } P = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

3. (a) True. In this case each eigenvalue is 1-fold. The matrix has a complete set of eigenvectors and is diagonalizable.

(b) False. It's possible for the 2-fold eigenvalue to have 2 corresponding ind eigenvectors in which case A is diagonalizable. Problem 1(d) is such an instance.

4. $A = P\Lambda_1 P^{-1}$ and $B = P\Lambda_2 P^{-1}$ so

$$AB = P\Lambda_1 P^{-1} \; P\Lambda_2 P^{-1} = P\Lambda_1 \Lambda_2 P^{-1}$$

$$BA = P\Lambda_2 P^{-1} \; P\Lambda_1 P^{-1} = P\Lambda_2 \Lambda_1 P^{-1}$$

Any two diagonal matrices commute (make up two 3×3 diagonal matrices and see that you get the same product if you multiply them in different orders). So $\Lambda_1 \Lambda_2 = \Lambda_2 \Lambda_1$ and so AB = BA.

5. (a) $P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ where $P = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \end{bmatrix}$

Also $P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ where $P = \begin{bmatrix} 5 & 1 & 9 \\ 6 & 2 & 10 \\ 7 & 3 & 11 \end{bmatrix}$

Also, you can use eigenvector 2u instead of u and get

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \text{ where } P = \begin{bmatrix} 2 & 5 & 9 \\ 4 & 6 & 10 \\ 6 & 7 & 11 \end{bmatrix} \quad \text{etc.}$$

(b) $|A|$ = product of the eigenvalues = 24

6. (a) Want to show that $(iK)^* = iK$.

$$
\begin{aligned}
(iK)^* &= \bar{i}\,K^* && (*\text{ rule}) \\
&= -iK^* \\
&= -i(-K) && (\text{since K is skew Herm}) \\
&= iK
\end{aligned}
$$

(b) iK is Herm so there exists a unitary U and real diagonal Λ so that $U^{-1}(iK)U = \Lambda$. So

$$(*) \qquad U^{-1}KU = \frac{1}{i}\Lambda = -i\Lambda$$

The matrix U is unitary. The matrix Λ is diagonal with real diagonal entries. So $-i\Lambda$ is diagonal with pure imaginary diagonal entries so (*) is the diagonalization of K we were looking for.

7. $A = P\Lambda P^{-1}$ so

$$
\begin{aligned}
A^* = (P\Lambda P^{-1})^* &= (P^{-1})^*\,\Lambda^*\,P^* && (*\text{ rules}) \\
&= (P^*)^*\,\Lambda^*\,P^{-1} && (\text{since P is unitary}) \\
&= P\Lambda^* P^{-1} \\
&= P\Lambda P^{-1} && (\Lambda^* = \Lambda \text{ because } \Lambda \text{ is real and diagonal}) \\
&= A
\end{aligned}
$$

So A is Hermitian.

8. $P^{-1}AP = \Lambda$ where $P = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\Lambda = \begin{bmatrix} .2 & 0 \\ 0 & 1 \end{bmatrix}$

So $A = P\Lambda P^{-1}$ and

$$
\begin{aligned}
A^\infty = P\Lambda^\infty P^{-1} &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} (.2)^\infty & 0 \\ 0 & 1^\infty \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} \\
&= \begin{bmatrix} 3 & -1 \\ 6 & -2 \end{bmatrix}
\end{aligned}
$$

SOLUTIONS review problems for Chapter 8

1. $|M - \lambda I| = \begin{vmatrix} 1-\lambda & 0 & 1 & 1 \\ 0 & 1-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 1 \\ 0 & 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^4$ So $\lambda=1$ is a 4-fold eigenvalue.

To find the corresponding eigenvectors solve

$$\left[\begin{array}{cccc|c} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Sol is $x_4 = 0$, $x_3 = -x_4 = 0$; x_1, x_2 free

Only 2 free variables so only 2 ind eigenvectors. M is not diagonalizable.

2. (a) Not necessarily.
   (b) Yes because (nonzero) multiples of eigenvectors are eigenvectors.

3. Given that there is some nonzero u such that $ABu = \lambda u$.
I want to find some nonzero vector v such that $BAv = \lambda v$.
Take the equation $ABu = \lambda u$ and left multiply by B (so that it at least begins with BA).

$$BABu = B\lambda u$$

Then rearrange to get

$$BA(Bu) = \lambda(Bu)$$

If $Bu \neq \vec{0}$ then Bu is an eigenvector of BA with corresponding eigenvalue λ.

Here's the proof (by contradiction) that Bu is not $\vec{0}$.

Suppose $Bu = \vec{0}$. Then $ABu = A\vec{0} = \vec{0}$, $\lambda u = \vec{0}$, $\lambda = 0$ or $u = \vec{0}$ which is impossible (we are given $\lambda \neq 0$ and we know $u \neq \vec{0}$ since it is an eigenvector). So $Bu \neq \vec{0}$. QED

4. (a) $Au = 3u$ so $u = \frac{1}{3}Au = A(\frac{1}{3}u)$. So the operator A sends $\frac{1}{3}u$ to u, i.e., u is the image of $\frac{1}{3}u$. So u is in the range of A.

(b) $Av = 0v = \vec{0}$. So the operator A sends v to $\vec{0}$. So v is in the null space of A.

5. Let $\Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$, $P = \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}$. Then choose $M = P\Lambda P^{-1} = \begin{bmatrix} 5 & -4/3 \\ 0 & 3 \end{bmatrix}$

6. (a) M can't be Herm because the eigenvalues aren't real.
   (b) M must be invertible because the eigenvalues are nonzero.
   (c) $|M|$ is the product of the eigenvalues. The det does not have mag 1. So M can't be unitary.

7. (a) Eigenvalues of A turn out to be 1,2 with corresponding eigenvectors (1,1), (2,1). One diagonalization is

$$P^{-1}AP = \Lambda \text{ where } P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

(b) $A = P\Lambda P^{-1}$ so

$$A^8 = P\Lambda^8 P^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2^8 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -1+2^9 & 2-2^9 \\ -1+2^8 & 2-2^8 \end{bmatrix}$$

8. For any matrix M, the product of the eigenvalues is $|M|$.
Suppose M is not invertible. Then $|M| = 0$.
So the product of the eigenvalues of M is 0.
That means one of the eigenvalues must be 0 (if the product is zero, one of the factors must be zero too).

9. Since A has eigenvalue 3, $\lambda=3$ is one of the roots of $|A - \lambda I| = 0$.
In other words, $|A - 3I| = 0$.
I want to show that 8A has eigenvalue 24 so I want to show that $\lambda=24$ is one of the roots of $|8A - \lambda I| = 0$. In other words, I want to show that $|8A - 24I| = 0$.

$$\begin{aligned} |8A - 24I| &= |8(A - 3I)| && \text{matrix algera} \\ &= 8^n \underbrace{|A - 3I|}_{0 \text{ by hypothesis}} && \text{det rule} \end{aligned}$$

So $|8A - 24I| = 0$, QED.

10. M can be diagonalized because it has enough eigenvectors so there exists a matrix P such that

$$P^{-1}MP = \underbrace{\begin{bmatrix} \text{1's and } -\text{1's} & 0 \\ 0 & \end{bmatrix}}_{\text{call this } \Lambda}$$

Then

$$M = P\Lambda P^{-1}$$

$$M^{-1} = (P\Lambda P^{-1})^{-1} = P\Lambda^{-1}P^{-1}$$

A diagonal matrix is inverted by just taking the reciprocal of each diagonal entry (the entry b becomes 1/b); 1's and −1's don't change when you do this so $\Lambda^{-1} = \Lambda$. So $M^{-1} = P\Lambda P^{-1} = M$.

SOLUTIONS Section 9.1

1. $A = \begin{bmatrix} 1 & -3/2 & 0 \\ -3/2 & 3 & -2 \\ 0 & -2 & 8 \end{bmatrix}$ and $q = \begin{bmatrix} x & y & z \end{bmatrix} A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

2. $q = 2x^2 + 6y^2 + 9z^2 + 6xy + 8xz + 14yz$

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

4. $x_1^2 + 5x_2^2 + 6x_4^2 + 4x_1x_2 + 6x_1x_3 + 8x_1x_4 - 6x_3x_4$

5. $P = \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix}$

(a) old matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, new matrix $= P^TAP = \begin{bmatrix} 24 & 1 \\ 1 & -1 \end{bmatrix}$

$q = 24X^2 + 2XY - Y^2$ in the new coord system

(b) $\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} X \\ Y \end{bmatrix}$

$x = 3X+2Y$, $y = X-Y$

$q = (3X+2Y)^2 + 4(3X+2Y)(X-Y) + 3(X-Y)^2 = 24X^2 + 2XY - Y^2$

(c) *Using X,Y coords.* Substitute $X=1, Y=2$ into the new q formula to get $q = 24+4-4 = 24$.
Using x,y coords. First get the x,y coords.

$P \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}$ so the old coords of B are $x = 7$, $y = -1$.

Substitute $x=7, y=-1$ into the old q formula to get $q = 49-28+3 = 24$ again.

6. (a) Solve for x and y to get $x = \frac{1}{2}(X + Y)$, $y = \frac{1}{2}(Y - X)$. Then substitute:

$$q = \frac{1}{4}(X + Y)^2 + 3\cdot\frac{1}{4}(X + Y)(Y - X) - 5\cdot\frac{1}{4}(Y - X)^2 = -\frac{7}{4}X^2 + 3XY - \frac{1}{4}Y^2$$

(b) Let P be the usual basis changing matrix (which converts from new to old coords). We know that

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{\text{this is } P^{-1}} \begin{bmatrix} x \\ y \end{bmatrix}$$

Take the inverse of P^{-1} to get $P = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$.

The new basis vectors are $u = (1/2, -1/2)$, $v = (1/2, 1/2)$.

(c) The (old) matrix for q is $A = \begin{bmatrix} 1 & 3/2 \\ 3/2 & -5 \end{bmatrix}$

Let P be the basis changing matrix from part (b) that converts x,y coords to X,Y coords.

The matrix for q in the new coord system is $P^T AP = \begin{bmatrix} -7/4 & 3/2 \\ 3/2 & -1/4 \end{bmatrix}$.

So $q = -\frac{7}{4} X^2 + 3XY - \frac{1}{4} Y^2$ again.

(d) The old matrix is $A = \begin{bmatrix} 1 & 3/2 \\ 3/2 & -5 \end{bmatrix}$; $q = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix}$

(e) The new matrix is B = X Y) $\begin{bmatrix} -7/4 & 3/2 \\ 3/2 & -1/4 \end{bmatrix}$; $q = \begin{bmatrix} X & Y \end{bmatrix} B \begin{bmatrix} X \\ Y \end{bmatrix}$

(f) $P^TAP = B$

7. (a) $q = (2X-Y)^2 + 4(2X-Y)(X+3Y) - (X+3Y)^2 = 11X^2 + 10XY - 20Y^2$

(b) Let P be the usual basis changing matrix (which converts from new to old coords). We know that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}}_{\substack{\text{this is the basis–} \\ \text{changing mattrix P}}} \begin{bmatrix} X \\ Y \end{bmatrix}$$

The new basis vectors are u = (2,1), v = (–1,3).

(c) The (old) matrix for q is $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$.

Let P be the basis changing matrix from part (b) that converts X,Y coords to x,y coords.
The matrix for q in the new coord system is

$$P^T AP = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 5 \\ 5 & -20 \end{bmatrix}$$

So $q = 11X^2 + 10XY - 20Y^2$ again.

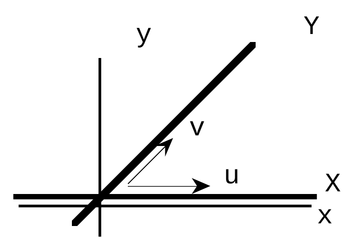
8. The new basis vectors u and v are unit vectors (because the scale in the new coord system is the same as the scale in the old system) pointing along the X-axis and Y-axis respectively:

$u = (1,0)$

$v = (1,1)_{\text{unit}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \left(\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}\right)$

Let

$$P = \begin{bmatrix} 1 & \sqrt{2}/2 \\ 0 & \sqrt{2}/2 \end{bmatrix}$$



Problem 8

(a) Old matrix $= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, new matrix $= P^T \text{old } P = \begin{bmatrix} 1 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1 \end{bmatrix}$

$q = X^2 + Y^2 + \sqrt{2}\,XY$

(b) New matrix for $q = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix} = P^T \text{ old } P$

$\text{old} = (P^T)^{-1} \text{ new } P^{-1} = (P^{-1})^T \text{ new } P^{-1} = \begin{bmatrix} 0 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2} \end{bmatrix}$

$q = -\sqrt{2}\, y^2 + \sqrt{2}\, xy$

(c) From part (a), $x^2 + y^2 = X^2 + Y^2 + \sqrt{2}\,XY$. So the circle has equation

$X^2 + Y^2 + \sqrt{2}\,XY = 1$ in the new coord system.

9. (a) *method 1*

When you change scales like this, the new coords X,Y and old coords x,y are related by $X = \frac{1}{12}x$, $Y = 2y$ (Section 24). So $x = 12X$, $y = \frac{1}{2}Y$ and

$$q = 2(12X)^2 + 3(12X)(\tfrac{1}{2}Y) + 4(\tfrac{1}{2}Y)^2 = 288X^2 + 18XY + Y^2$$

method 2

The new basis vectors are $u = 12i$, $v = \frac{1}{2}j$. So $P = \begin{bmatrix} 12 & 0 \\ 0 & 1/2 \end{bmatrix}$

New matrix $= P^T \begin{bmatrix} 2 & 3/2 \\ 3/2 & 4 \end{bmatrix} P = \begin{bmatrix} 288 & 9 \\ 9 & 1 \end{bmatrix}$ so $q = 288X^2 + 18XY + Y^2$

SOLUTIONS Section 9.2

1. (a) (i) q has matrix A = $\begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.

The eigenvalues of A are 2,−3 so $q = 2X^2 - 3Y^2$.
The corresponding eigenvectors are (2,1), (−1,2).
Ortho*normal* eigenvectors are $u = (2/\sqrt{5},\ 1/\sqrt{5})$, $v = (-1/\sqrt{5},\ 2/\sqrt{5})$.
The coord system in which q is $2X^2 - 3Y^2$ has basis vectors u,v.
v

(ii) $\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} X \\ Y \end{bmatrix}$ where $P = \begin{bmatrix} u & v \end{bmatrix}$ so the change of variable is

$$x = \frac{2}{\sqrt{5}} X - \frac{1}{\sqrt{5}} Y,\quad y = \frac{1}{\sqrt{5}} X + \frac{2}{\sqrt{5}} Y$$

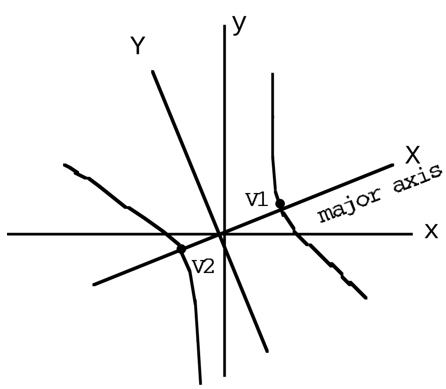
(iii) The old matrix for q was found in part (i); $q = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix}$

(iv) The new matrix for q is $B = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$; $q = \begin{bmatrix} X & Y \end{bmatrix} B \begin{bmatrix} X \\ Y \end{bmatrix}$

(v) $B = P^T A P$ where $P = \begin{bmatrix} u & v \end{bmatrix}$.

(b) Use the orthonormal coord system from part (a) with basis u,v. In the new system, the equation is $2X^2 - 3Y^2 = 2$.

The graph is a hyperbola whose major axis is the X−axis, the line $y = \frac{1}{2}x$. The vertices are ((±1,0)) which in the original coord system are $(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})$ and $(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}})$.



Problem 1(b)

2.(a) Let $q = 2x^2 + 2y^2 + 3z^2 + 4xy + 2xz + 2yz$. Matrix for q is $\begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}$.

Eigenvalues are 2,5,0 with corresponding eigenvectors
u = (−1,−1,2), v = (1,1,1), w = (1,−1,0)

They are already orthog. Normalize them to get orthonormal basis vectors. In this new orthonormal coord system, $q = 2X^2 + 5Y^2$ so the equation is $2X^2 + 5Y^2 = 9$
It's an elliptic cylinder with its axis along the Z-axis (see the diagram).
(b) If you use another method of diagonalizing, the new coord system won't necessarily be orthonormal. And if it isn't then you can't tell a circular cylinder from an elliptical cylinder (great tragedy).

(c) Let $P = [u_{unit}\ v_{unit}\ w_{unit}] = \begin{bmatrix} -1\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ -1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \\ 2/\sqrt{6} & 1/\sqrt{3} & 0 \end{bmatrix}$

Then $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ so

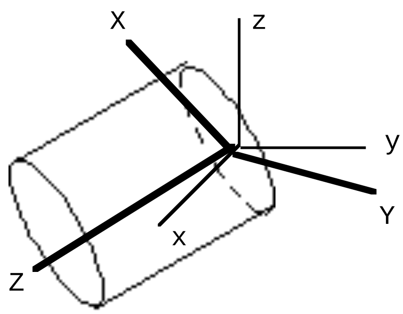
$x = -\frac{1}{\sqrt{6}} X + \frac{1}{\sqrt{3}} Y + \frac{1}{\sqrt{2}} Z$, $y = -\frac{1}{\sqrt{6}} X + \frac{1}{\sqrt{3}} Y - \frac{1}{\sqrt{2}} Z$, $z = \frac{2}{\sqrt{6}} X + \frac{1}{\sqrt{3}} Y$.

To solve for X,Y,Z in terms of x,y,z take advantage of the fact that $P^{-1} = P^T$ since P is an orthog matrix. So

$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = P^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = P^T \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$X = -\frac{1}{\sqrt{6}} x - \frac{1}{\sqrt{6}} y + \frac{2}{\sqrt{6}} z$, $Y = \frac{1}{\sqrt{3}} x + \frac{1}{\sqrt{3}} y + \frac{1}{\sqrt{3}} z$, $Z = \frac{1}{\sqrt{2}} x - \frac{1}{\sqrt{2}} y$

(d) Same q as in part (b) so the equation is $2X^2 + 5Y^2 = -9$. Graph is empty (no points satisfy the equation).



Problem 2(a)

3. (a) Let $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.

Do ops $R_2 = -2R_1 + R_2$, $C_2 = -2C_1 + C_2$ on A to get $\begin{bmatrix} 1 & 0 \\ 0 & -6 \end{bmatrix}$.

Do the col op on I to get $P = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$.

Then $q = X^2 - 6Y^2$ in the coord system with basis vectors u = (1,0), v − (−2,1).

(b) $\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$ so $x = X - 2Y$, $y = Y$.

The check i sthat $(X-2Y)^2 + 4(X-2Y)Y - 2Y^2$ does equal $X^2 - 6Y^2$.

(c) $P^TAP = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -6 \end{bmatrix}$ which *is* what I got with the row/col op method.

4. Do these row/col ops on A.

$R_2 = -R_1 + R_2$, $C_2 = -C_1 + C_2$

$R_3 = -2R_1 + R_3$, $C_3 = -2C_1 + C_3$

$R_3 = -\frac{3}{2}R_2 + R_3$, $C_3 = -\frac{3}{2}C_2 + C_3$

and get $\begin{matrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1/2 \end{matrix}$

This is enough to diagonalize q but if you want ±1's keep going.
Divide Row 2 by $\sqrt{2}$, divide Col 2 by $\sqrt{2}$. Multiply Row 3 by $\sqrt{2}$, multiply Col 3 by $\sqrt{2}$ and get $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Do all the col ops to I to get $P = \begin{bmatrix} 1 & -1\sqrt{2} & -1/\sqrt{2} \\ 0 & 1\sqrt{2} & -3/\sqrt{2} \\ 0 & 0 & \sqrt{2} \end{bmatrix}$

In the coord system with the cols of P as the basis vectors, $q = X^2 + Y^2 + Z^2$.

(b) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ so $x = X - \frac{1}{\sqrt{2}} Y - \frac{1}{\sqrt{2}} Z$, $y = \frac{1}{\sqrt{2}} Y - \frac{3}{\sqrt{2}} Z$, $z = \sqrt{2}\, Z$

5. Use row/col ops to diagonalize.
$R_3 = -R_1 + R_3$; $C_3 = -C_1 + C_3$
$R_3 = -3R_2 + R_3$; $C_3 = -3C_2 + C_3$
Matrix becomes $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -7 \end{bmatrix}$.
The diagonal has 2 positive and 1 negative entry.
So there are 2 positive eigenvalues and one negative eigenvalue.

6. First, use the given information to get a diagonal version of q.
$q = \frac{1}{3}X^2 - 4Y^2$ in a coord system with basis $u = (-\frac{4}{5}, \frac{3}{5})$, $v = (\frac{3}{5}, \frac{4}{5})$
(Remember to normalize the orthogonal eigenvectors.)

Now that q is diagonal, here are two ways to get the diagonal coeffs to be ±1's.
method 1 (row/col ops) The matrix for q w.r.t. basis u,v is $A = \begin{bmatrix} 1/3 & 0 \\ 0 & -4 \end{bmatrix}$.
Do the row/col ops

multiply row 1 by $\sqrt{3}$; mult col 1 by $\sqrt{3}$
divide row 2 by 2; divide col 2 by 2.

on A to get $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

Let $P = \begin{bmatrix} -4/5 & 3/5 \\ 3/5 & 4/5 \end{bmatrix}$. Do the col ops on P to get $Q = \begin{bmatrix} -4\sqrt{3}/5 & 3/10 \\ 3\sqrt{3}/5 & 4/10 \end{bmatrix}$.

Then $q = X^2 - Y^2$ in a coord system with basis vectors
$u_1 = (-4\sqrt{3}/5,\ 3\sqrt{3}/5)$, $v_1 = (3/10,\ 4/10)$ (the cols of Q).

method 2 for continuing (mostly by inspection and a little algebra)
$q = \frac{1}{3}X^2 - 4Y^2$ in a coord system with basis $u = (-\frac{4}{5}, \frac{3}{5})$, $v = (\frac{3}{5}, \frac{4}{5})$.

q can be rewritten as $(\frac{1}{\sqrt{3}} X)^2 - (2Y)^2$.

Let $X_1 = \frac{1}{\sqrt{3}} X$, $Y_1 = 2Y$. Then $q = X_1^2 - Y_1^2$ in a new X_1, Y_1 coord system.

Now I just have to get the new basis vectors u_1 and v_1.

The change of variable from X,Y to X_1, Y_1 just changed the scales on the X-axis and Y-axis (see (5), (6), (7) in Section 2.4):

$$X_1\text{-scale} = \sqrt{3}\ X\text{-scale}$$

$$Y_1\text{-scale} = \frac{1}{2}\ Y\text{-scale}$$

$$u_1 = \sqrt{3}\ u = (-\ 4\sqrt{3}/5,\ 3\sqrt{3}/5)$$

$$v_1 = \frac{1}{2}\ v = (3/10,\ 4/10)$$

7. (a) $q = (x^2 + 4xy + 4y^2) - 2y^2 - 4y^2$

$$= (x+2y)^2 - 6y^2$$

$$= X^2 - 6Y^2$$

where

$$X = x + 2y, \quad Y = y$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}}_{P^{-1}} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

The new basis vectors are u = (1,0), v = (−2,1)

(b) $q = 2(x^2 + \frac{3}{2}xy\quad) + y^2$

$$= 2(x^2 + \frac{3}{2}xy + \frac{9}{16}y^2) + y^2 - \frac{9}{8}y^2$$

$$= 2(x + \frac{3}{4}y)^2 - \frac{1}{8}y^2$$

$$= 2X^2 - \frac{1}{8}Y^2$$

where

$$X = x + \frac{3}{4}y, \quad Y = y$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 3/4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$P^{-1} = \begin{bmatrix} 1 & 3/4 \\ 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & -3/4 \\ 0 & 1 \end{bmatrix}$$

The new basis vectors are u = (1,0), v = (−3/4, 1)

(c) $q = 3\left[x^2 + (2y+6z)x\right] - 6y^2 + z^2$

The coeff of x in the brackets is 2y+6z. Take half, square it and add it on to complete the square.

$$q = 3\left[x^2 + (y+3z)\right]^2 - 6y^2 + z^2 - 3(y+3z)^2$$

$$= 3\left[x + y + 3z\right]^2 - 9y^2 - 18yz - 26z^2$$

$$= 3\left[x + y + 3z\right]^2 - 9\left[y^2 + 2zy\right] - 26z^2$$

The coeff of y in the second bracket is 2z. Take half, square it and add it on to complete the square.

$$q = 3\left[\, x + y + 3z \,\right]^2 - 9\left[\, y^2 + 2zy + z^2 \,\right] - 26z^2 + 9z^2$$
$$= 3\left[\, x + y + 3z \,\right]^2 - 9\left[\, y + z \,\right]^2 - 26z^2 + 9z^2 - 17z^2$$
$$= 3X^2 - 9Y^2 - 17Z^2$$

where

$$X = x + y + 3z,\ Y = y + z\ ,\ Z = z\ .$$

Then

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$P^{-1} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$ (I used Mathematica to invert)

so the new basis vectors are $u = (1,0,0)$, $v = (-1,1,0)$, $w = (-2,-1,1)$.

(d) $q = 3\left[x^2 + \frac{2y+5z}{3}\, x\right] + y^2 + 4z^2 + 6yz$

In the bracket, the coeff of x is $\frac{2y+5z}{3}$. Take half, square it and add it on to complete the square.

$$q = 3\left[\, x + \frac{2y+5z}{6} \,\right]^2 + y^2 + 4z^2 + 6yz - 3\left(\frac{2y+5z}{6}\right)^2$$
$$= 3\left[\, x + \frac{2y+5z}{6} \,\right]^2 + \frac{2}{3}y^2 + \frac{13}{3}yz + \frac{23}{12}z^2$$
$$= 3\left[\, x + \frac{2y+5z}{6} \,\right]^2 + \frac{2}{3}\left[y^2 + \frac{13}{2}z\,y\right] + \frac{23}{12}z^2$$

In the second bracket, the coeff of y is $\frac{13}{2}$ z. Take half, square it and add it on to complete the square.

$$q = 3\left[\, x + \frac{2y+5z}{6} \,\right]^2 + \frac{2}{3}\left[y + \frac{13}{4}z\right]^2 + \frac{23}{12}z^2 - \frac{2}{3}\left(\frac{13}{4}z\right)^2$$
$$= 3\left[\, x + \frac{2y+5z}{6} \,\right]^2 + \frac{2}{3}\left[y + \frac{13}{4}z\right]^2 - \frac{41}{8}z^2$$

Let $X = x + \frac{1}{3}y + \frac{5}{6}z$, $Y = y + \frac{13}{4}z$, $Z = z$. Then $q = 3X^2 + \frac{2}{3}Y^2 - \frac{41}{8}Z^2$.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1/3 & 5/6 \\ 0 & 1 & 13/4 \\ 0 & 0 & 1 \end{bmatrix}}_{P^{-1}} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Invert P^{-1} to get P. The new basis vectors are the cols of P.

SOLUTIONS Section 9.3

1. *method 1* Do these row/col ops

$$R_2 = -\frac{1}{2}R_1 + R_2, \quad C_2 = -\frac{1}{2}C_1 + C_2$$

$$R_3 = -\frac{1}{2}R_1 + R_3, \quad C_3 = -\frac{1}{2}C_1 + C_3$$

$$R_3 = R_2 + R_3, \; C_3 = C_2 + C_3$$

A turns into $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. The diagonal entries are ≥ 0 so A is positive semi-definite (and one diagonal entry actually *is* 0 so A is not positive definite).

method 2 The ℓpm's are $\det[2] = 2$, $\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1$, $|A| = 0$.

The ℓpm's are ≥ 0 so A is positive semi-def (and not positive definite).

.

method 3

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 1-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)(1-\lambda) - (1-\lambda) - (1-\lambda) = (1-\lambda)(\lambda^2 - 3\lambda)$$

Eigenvalues are 1,0,3, all ≥ 0 and one actually *is* 0 so A is positive semi-def (and not positive definite).

2. Many methods available. For 2×2 matrices using ℓpm's is fastest.

(a) $\det[2] = 2$, $\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3$. The ℓpm's are positive so matrix is pos def.

(b) $\det[1] = 1$, $\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3$. Matrix is indefinite.

3. False (it's true for a diagonalized q, with only square terms, but not necessarily true otherwise). As a counterexample, let $q = 2x^2 + 4xy + y^2$. Then q is negative when $x=-1$, $y=1$ so q is not positive definite.

4. (a) Let q be the quadratic form with matrix A. Then $-A$ has quadratic form $-q$.

For example, if $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ then the quadratic form with matrix A is $ax^2 + 2bxy + cy^2$ and the quadratic form with matrix $-A$ is $-ax^2 - 2bxy - cy^2$.

If A is neg definite then $q < 0$ (except at the origin) so $-q > 0$ (except at the origin) so $-A$ is positive definite.

(b) The ℓpm's of A have signs $- + - + - +$ etc.

What happens to the ℓpm's when you multiply A by -1.

Remember that when you multiply an $n \times n$ matrix by -1, the determinant is multiplied by $(-1)^n$. That means it stays the same if n is even and changes sign if n is odd.

So the first ℓpm , a 1×1 det which was negative, changes sign.
The second ℓpm, a 2×2 det which was positive, doesn't change sign.
The third ℓpm, a 3×3 det which was negative, changes sign. And so on.
So the ℓpm's of $-A$ are all positive. So $-A$ is positive definite.

5. *method 1* Let q be the quadratic form corresponding to A.
Then 3A corresponds to quadratic form 3q.
If $q > 0$ (except at the origin) then $3q > 0$ (except at the origin)
So if A is positive definite then 3A is also positive definite.

method 2
The ℓpm's of A are positive
When you multiply A by 3, you multiply the first ℓpm by 3, the second ℓpm by 3^2, the third ℓpm by 3^3 etc. The new ℓpm's are still positive so 3A is positive definite too.

6. (a) A is negative definite (the diagonal form of q has all negative coeffs).
So the eigenvalues are negative (not necessarily the numbers −2,−3,−5, but all negative).
$|A|$ is negative.
reason 1 it's the product of the 3 negative eigenvalues
reason 2 $|A|$ is the 3rd ℓpm and the signs of the ℓpm's are − + −
(b) A is indefinite;
A has two positive eigenvalues and one negative eigenvalue;
$|A|$ is the product of the eigenvalues so it's negative.

7. (a) $(B^TB)^T = B^TB^{TT} = B^TB$. So B^TB is symmetric.
(b) Let $\vec{x}$ be the col vector of variables. Then $q = x^T(B^TB)x$
(c) $q = x^T(B^TB)x = (Bx)^T Bx$ (T rule)
$= Bx \cdot Bx$ (connection between dotting vectors and multiplying matrices)
$= \|Bx\|^2$ (connection between norms and dots)
≥ 0 (since norms are never negative)

So q is positive semi-definite.

8. (a) True. *reason 1* $|A|$ is one of the ℓpm's and all the ℓpm's are positive.
reason 2 All the eigenvalues of A are positive and $|A|$ is the product of the eigenvalues.
(b) False. A counterexample is $A = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}$.
$|A|$ is positive but A is negative definite (its ℓpm's are −2, +6).

9. The matrix for q is $\begin{bmatrix} 2 & b/2 \\ b/2 & 3 \end{bmatrix}$. The ℓpm's are 2 and $6 - \frac{1}{4}b^2$.
(a) q is positive def iff $6 - \frac{1}{4}b^2 > 0$, $-\sqrt{24} < b < \sqrt{24}$
(b) q is positive semi-def iff $6 - \frac{1}{4}b^2 \geq 0$, $-\sqrt{24} \leq b \leq \sqrt{24}$
(c) q is positive semi-def but not pos def iff $b = \pm\sqrt{24}$
(d) q is indefinite iff $6 - \frac{1}{4}b^2 < 0$, $b > \sqrt{24}$ or $b < -\sqrt{24}$
(e) Can never have q negative definite.

10. (a) *proof 1*
Step 1 The eigenvalues of A are nonzero. That's because a positive definite matrix always has positive eigenvalues which means that as a by-product, the eigenvalues are nonzero.

Step 2 A matrix with nonzero eigenvalues is invertible (see the latest invertible list in Section 8.2).

proof 2 The ℓpm's of A are positive. But the last ℓpm is $|A|$ itself. So $|A| \neq 0$ so A is invertible.
(b) The converse (that if A is invertible then it is positive definite) is false. As a counterexample, the matrix −I is invertible but isn't positive def (it's neg def).

(c) First I'll show that the eigenvalues of A^{-1} and A are reciprocals of one another.
Suppose λ is an eigenvalue of A with corresponding eigenvector u.
Then $Au = \lambda u$ where $u \neq \vec{0}$ so $u = A^{-1}\lambda u = \lambda A^{-1} u$.
$\lambda \neq 0$ since A is invertible so it's safe to divide by λ to get $A^{-1} u = \frac{1}{\lambda} u$.

This makes u an eigenvector of A^{-1} with corresponding eigenvalue $1/\lambda$.
So A and A^{-1} have reciprocal eigenvalues.

So if the eigenvalues of A are positive then the eigenvalues of A^{-1} must also be positive. So if A is positive definite then so is A^{-1} QED

11. (a) True. Here's a proof by contradiction.
The quadratic form with matrix A is $q = ax^2 + dy^2 + fz^2 + 2bxy + 2cxz + 2eyz$.
Suppose d was negative. Then q would be negative when x=0,y=1, z=0 which contradicts the fact that A is positive definite.
Suppose d was zero. Then q = 0 when x=0, y=0, z=1 which again contradicts the fact that A is positive definite.
So d must be positive. Similarly, a and f must be positive.

(b) False. Here's a counterexample. Let $A = \begin{bmatrix} 1 & 2 & \cdot \\ 2 & 1 & \cdot \\ \cdot & \cdot & 1 \end{bmatrix}$. The diagonal entries of A are positive but A is not positive definite. One reason is that the second ℓpm is not positive. Another reason is that the quadratic form with matrix A is
$q = x^2 + y^2 + z^2 + 4xy$ + who cares. It is negative when $x=1, y=-1, z=0$.

SOLUTIONS review problems for Chapter 9

1. q has matrix $A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$.

(a) (i) The new basis just amounts to a change of scale on the old axes (see "special case where only the scale changes" in §2.4).

So $X = \frac{1}{2}x$, $Y = 3y$; $x = 2X$, $y = \frac{1}{3}y$. So

$q = 2x^2 - 4xy + 5y^2 = 2(2X)^2 - 4(2X)(\frac{1}{3}Y) + 5(\frac{1}{3}Y)^2 = 8X^2 - \frac{8}{3}XY + \frac{5}{9}Y^2$

(ii) Let $P = \begin{bmatrix} 2 & 0 \\ 0 & 1/3 \end{bmatrix}$. Then

new matrix for q is $P^TAP = \begin{bmatrix} 8 & -4/3 \\ -4/3 & 5/9 \end{bmatrix}$. So $q = 8X^2 - \frac{8}{3}XY + \frac{5}{9}Y^2$.

(b) *method 1* Eigenvalues are A are 6,1 with corresponding eigenvectors (1,−2), (2,1). They are already orthogonal. Normalize them to get

$u = (\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}})$, $v = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})$

With basis u,v, $q = 6X^2 + Y^2$.

If you use the eigenvectors (1,−2) and (2,1) as the basis, without normalizing them, then q is still diagonal but it's $30X^2 + 5Y^2$, not $6X^2 + Y^2$.

method 2 Do row/col op $R_2 = R_1 + R_2$, $C_2 = C_1 + C_2$. A turns into $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

Do the col op to I and get $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. With new basis u = (1,0), v = (1,1) (*don't* normalize) the quadratic form is $q = 2X^2 + 3Y^2$.

If you want to normalize u and v (if you have some desperate compulsion to normalize) then you can use new basis vectors u_{unit}, v_{unit} but then q is $2X^2 + \frac{3}{2}Y^2$, not $2X^2 + 3Y^2$.

method 3

$$\begin{aligned} q &= 2(x^2 - 2xy \quad) + 5y^2 \\ &= 2(x^2 - 2xy + y^2) + 5y^2 - 2y^2 \\ &= 2(x-y)^2 + 3y^2 \\ &= 2X^2 + 3Y^2 \end{aligned}$$

where

$$X = x-y, \quad Y = y.$$

Then

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

so

$$P^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The new basis vectors are u = (1,0), v = (1,1).

(c) The best system in which to graph is the orthonormal system from method 1. The equation is $6X^2 + Y^2 = 7$ and because it's an orthonormal system you know that the graph is an ellipse (not a circle).

The major axis is the Y-axis, the line through the origin pointing like vector (2,1). So the major axis is line $y = \frac{1}{2}x$ (see the diagram).

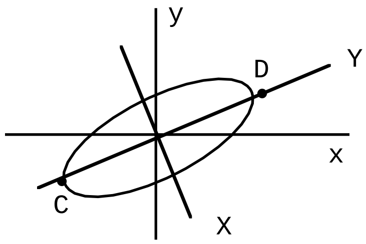
The minor axis is the X-axis, the line through the origin pointing like vector (1,−2). So the minor axis is the line $y = -2x$.

The vertices of the ellipse are points C and D. In the new coord system, D has coords X=0, $Y = \sqrt{7}$.

$$P\begin{bmatrix} 0 \\ \sqrt{7} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 0 \\ \sqrt{7} \end{bmatrix} = \begin{bmatrix} 2\sqrt{7}/\sqrt{5} \\ \sqrt{7}/\sqrt{5} \end{bmatrix}$$

So the old coords of D are $x = 2\sqrt{7}/\sqrt{5}$, $y = \sqrt{7}/\sqrt{5}$.

Similarly, C has new coords $X = 0$, $Y = -\sqrt{7}$ and old coords $x = -\dfrac{2\sqrt{7}}{\sqrt{5}}$, $y = -\dfrac{\sqrt{7}}{\sqrt{5}}$.



Problem 1(c)

2. (a) *method 1* The ℓpm's are 2,3,4 so A is positive definite.

method 2 Use row/col ops

add $\frac{1}{2}$ row 1 to row 2, add $\frac{1}{2}$ col 1 to col 2

add $-\frac{1}{2}$ row 1 to row 3, add $-\frac{1}{2}$ col 1 to col 3

add $\frac{1}{3}$ row 2 to row 3, add $\frac{1}{3}$ col 2 to col 3

New matrix is $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 4/3 \end{bmatrix}$. Diagonal entries are positive so A is pos def.

(b) Continue from the row/col ops in part (a).

divide row 1 by $\sqrt{2}$, divide col 1 by $\sqrt{2}$

multiply row 2 by $\sqrt{2/3}$, multiply col 2 by $\sqrt{2/3}$

multiply row 3 by $\sqrt{3/4}$, multiply col 3 by $\sqrt{3/4}$

All the row/col ops put together turn A into I.

Do all the col ops to I to get $P = \begin{bmatrix} 1/\sqrt{2} & \sqrt{2}/(2\sqrt{3}) & -\sqrt{3}/6 \\ 0 & \sqrt{2}/3 & \sqrt{3}/6 \\ 0 & 0 & \sqrt{3}/2 \end{bmatrix}$

Then $P^TAP = I$

3. $q = \vec{x}^{\,T} A \vec{x}$

4. There is some (invertible) P so that $A = P^TBP$. Then

$$\begin{aligned} |AB| &= |P^TBPB| \\ &= |P^T|\ |B|\ |P|\ |B| && \text{(det rule)} \\ &= |P|^2\ |B|^2 && (|P^T| = |P|\ ) \\ &\ge 0 && \text{(the product of real squares is} \ge 0) \end{aligned}$$

(Didn't need the fact that P is invertible.)

5. $M^TAM = B$ (The quadratic form that has matrix A w.r.t. the standard basis has matrix B w.r.t. the new basis composed of the cols of M.)

6. old = $\begin{bmatrix} 1 & 3/2 \\ 3/2 & -1 \end{bmatrix}$, new = $\begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$.

Let P have cols u and v. Then P^T old P = new.

SOLUTIONS Section 10.1

1. $(au)\cdot v = \overline{v\cdot(au)}$ by dot axiom 1

$= \overline{a(v\cdot u)}$ by dot axiom 3

$= \overline{a}\;\overline{v\cdot u}$ by the algebra rule $\overline{zw} = \overline{z}\,\overline{w}$ for complex numbers

$= \overline{a}\,(u\cdot v)$ by dot axiom 1 again

2. I'm right. His operation satisfies dot axioms 1–4 so it's a legal dot.

The student's vectors u and v are orthogonal w.r.t. the standard dot product in R^2 but that doesn't mean that they have to be orthogonal w.r.t. another dot product. Orthogonality is relative. It depends on the dot.

3. (a) Not a dot. Here's one reason why not.

If $u = (2,-3)$ then $u \bullet u = -12$. But this violates the dot rule $u\cdot u \ge 0$.

(b) Can't possibly be a dot product since this operation produces a vector, not scalar.

(c) Yes. It satisfies dot axioms 1–4.

4. (a)(i) A natural basis is

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},\ A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},\ A_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Every 2×3 matrix can be written in exactly one way in terms of them.

For instance, if $M = \begin{bmatrix} 5 & 6 & \pi \\ \sqrt{2} & 7 & 8 \end{bmatrix}$ then $M = 5A_1 + 6A_2 + \pi A_3 + \sqrt{2}A_4 + 7A_5 + 8A_6$.

The set of 2×3 matrices is a 6-dim vector space.

(b) If A and B are symmetric then A + B and kA are also symmetric (this was checked out in a problem in Section 1.6). So the set of symmetric 3×3 matrices is closed under addition and scalar mult so it's a subspace.

One basis is

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ B_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ B_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B_4 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ B_5 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},\ B_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

For instance, if $M = \begin{bmatrix} 3 & 4 & 5 \\ 4 & 6 & 7 \\ 5 & 7 & 8 \end{bmatrix}$ then $M = 3B_1 + 6B_2 + 8B_3 + 4B_4 + 5B_5 + 7B_6$

So the subspace is 6-dim. (The whole space of all 3×3 matrices is 9-dim.)

(c) No. *reason 1* The set of invertible matrices doesn't include the zero matrix.

reason 2 The set of invertibles isn't closed under addition. If

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

then A and B are invertible, but A + B isn't.

5. (a) Need to show that dot axioms 1 (real version), 2, 3, 4 hold.

axiom 1 (real version) $y \circ x = y_1 x_1 - y_1 x_2 - y_2 x_1 + 2y_2 x_2$

$x \circ y$ comes out to be the same thing so $y \circ x = x \circ y$

axiom 2

$$x \circ x = x_1 x_1 - x_2 x_1 - x_1 x_2 + 2x_2 x_2 = x_1^2 - 2x_1 x_2 + 2x_2^2 = (x_1 - x_2)^2 + x_2^2$$

So $x \circ x$ is the sum of (real) squares so it's ≥ 0, and it *equals* 0 iff both x_1 and x_2 are 0, i.e., iff $x = \vec{0}$.

axiom 3 $x \circ (ay) = x_1(ay_1) - x_2(ay_1) - x_1(ay_2) + 2x_2(ay_2)$
$= a(x_1 y_1 - x_2 y_1 - x_1 y_2 + 2x_2 y_2) = a(x \circ y)$

axiom 4 (boring)

$x \circ (u + v) = x_1(u_1 + v_1 - x_2(u_1 + v_1) - x_1(u_2 + v_2) + 2x_2(u_2 + v_2)$

$x \circ u + x \circ v = x_1 u_1 - x_2 u_1 - x_1 u_2 + 2x_2 u_2 + x_1 v_1 - x_2 v_1 - x_1 y_2 + 2x_2 v_2$

Just look to see that $x \circ (u + v)$ equals $x \circ u + x \circ v$

(b) $i \circ j = (1)(0) - (0)(0) - (1)(1) + 2(0)(1) = -1$ Not orthog w.r.t. new dot.

(c) $u \circ v = 2 \times 4 - 3 \times 4 - 2 \times 1 + 2 \times 3 \times 1 = 0$ Orthog w.r.t. new dot.

6. (a) The natural basis is

$p_4 = x^4,\ p_3 = x^3,\ p_2 = x^2,\ p_1 = x,\ p_0 = 1$

Then $p = 5p_4 + 0p_3 + 2p_2 - p_1 + 2p_0$ so p has coords 5,0,2,−1,2 w.r.t. the standard basis.

The vector space has degree 5.

(b) First write the new basis vectors in terms of the old ones.

$$\begin{aligned}
q_4 &= 3x^4 &&= 3p_4 + 0p_3 + 0p_2 + 0p_1 + 0p_0\\
q_3 &= 2x^3 &&= 0p_4 + 2p_3 + 0p_2 + 0p_1 + 0p_0\\
q_2 &= (x-3)^2 &&= 0p_4 + 0p_3 + 1p_2 - 6p_1 + 9p_0\\
q_1 &= x + 5 &&= 0p_4 + 0p_3 + 0p_2 + 1p_1 + 5p_0\\
q_0 &= 6 &&= 0p_4 + 0p_3 + 0p_2 + 0p_1 + 6p_0
\end{aligned}$$

Let P be the matrix whose cols are the coordinates of the new basis vectors $q_4,\ldots,q_0$ w.r.t. the standard basis vectors $p_4,\ldots,p_0$:

$$P = \begin{bmatrix} 3 & 0 & 0 & 0 & 0\\ 0 & 2 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & -6 & 1 & 0\\ 0 & 0 & 9 & 5 & 6 \end{bmatrix}$$

Then

$$P^{-1} = \begin{bmatrix} 1/3 & 0 & 0 & 0 & 0\\ 0 & 1/2 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 6 & 1 & 0\\ 0 & 0 & -13/2 & -5/6 & 1/6 \end{bmatrix}, \quad P^{-1}\begin{bmatrix} 5\\ 0\\ 2\\ -1\\ 2 \end{bmatrix} = \begin{bmatrix} 5/3\\ 0\\ 2\\ 11\\ -71/6 \end{bmatrix}$$

So $p = \frac{5}{3} q_4 + 2q_2 + 11q_1 - \frac{71}{6} q_0.$

Check If you take $\frac{5}{3} 3x^4 + 2(x-3)^2 + 11(x+5) - \frac{71}{6} 6$ and multiply it all out you do get $5x^4 + 2x^2 - x + 2$.

(c) The set of polys of degree 4 isn't closed under addition. If you add the polys $2x^4 + x^3$ and $-2x^4$ you get x^3 which is not a poly of degree 4.

SOLUTIONS Section 10.2

1. $< f, g > = \int_1^2 2x^3\, dx = \frac{30}{4}$

2.(a) $< f, g > = \int_0^1 \overline{2ix}\ (3x + ix)\, dx = \int_0^1 -2ix\,(3x + ix)\, dx$

$= \int_0^1 2x\, dx - i \int_0^1 6x^2\, dx = \frac{2}{3} - 2i$

$< g, f > = \overline{< f, g >} = \frac{2}{3} + 2i$

(b) $< f, g > = \int_{-1}^1 2x\, dx - i \int_{-1}^1 6x^2\, dx = -4i$

$< g, f > = 4i$

3. $< 1, \cos \frac{7\pi x}{L} > = \int_0^L \cos \frac{7\pi x}{L}\, dx = \frac{L}{7\pi} \sin \frac{7\pi x}{L} \Big|_0^L = 0$

so 1 and $\cos \frac{7\pi x}{L}$ are orthog on $[0,L]$.

4. (a) On interval $[0,2\pi]$,

$< f(x), g(x) > = \int_{x=0}^{2\pi} (\cos x - i \sin x)(2 + 3i)\, dx$

$= \int_{x=0}^{2\pi} (2 \cos x + 3 \sin x)\, dx + i \int_{x=0}^{2\pi} (3 \cos x - 2 \sin x)\, dx$

$= 0$ because $\int_0^{2\pi} \cos x\, dx = \int_0^{2\pi} \sin x\, dx = 0$.

So f and g are orthogonal on $[0, 2\pi]$.

(b) On interval $[0,\pi]$,

$< f(x), g(x) > = \int_{x=0}^{\pi} (2 \cos x + 3 \sin x)\, dx + i \int_{x=0}^{\pi} (3 \cos x - 2 \sin x)\, dx$

$= 6 - 4i$ because $\int_0^{\pi} \cos x\, dx = 0$, $\int_0^{\pi} \sin x\, dx = 2$

So f and g are not orthogonal on $[0,\pi]$.

5. $< f, f' > = \int_a^b f(x)\, f'(x)\, dx = \frac{1}{2} f^2(x) \Big|_a^b = \frac{1}{2} f^2(b) - \frac{1}{2} f^2(a) = 0$

so f and f' are orthog on $[a,b]$.

6. (a) *axiom 1* $f \bullet g = \int_a^b f(x)g(x)\, x^2\, dx$

$g \bullet f = \int_a^b g(x)f(x)\, x^2\, dx$

The integrals are the same so $f \bullet g = g \bullet f$.

axiom 2 $\quad f \bullet f = \int_a^b f^2(x)\ x^2\ dx$

The function $x^2\ f^2(x)$ is always ≥ 0. So the integral is always ≥ 0. Furthermore, the integral can *be* zero iff $f^2(x)$ is the zero function.

(b) Axiom 2 is violated because $f \bullet f$ is not always ≥ 0. For instance, if $f(x) = x^2$ then

$$f \bullet f = \int_{-1}^{0} (x^2)^2\ x^3\ dx = \int_{-1}^{0} x^7\ dx = -\frac{1}{8}$$

7. The vectors x and 1 for instance are not orthogonal on [0,1] because

$$< x,\ 1 > = \int_0^1 x\ dx \neq 0$$

So the natural basis is not orthogonal on [0,1].
Use the Gram Schmidt process to get an orthogonal basis $u_1, \ldots, u_5$.

I'll use notation $x_1, \ldots, x_5$ for the standard basis x^4, x^3, x^2 x, 1 so that my notation matches the Gram Schmidt notation. (Awkward notation: the function x^3 is given the vector name $\vec{x}_2$). Then

$$u_1 = x_1 = x^4$$

$$u_2 = x_2 - \frac{u_1 \cdot x_2}{u_1 \cdot u_1} u_1 = x^3 - \frac{\int_0^1 x^7\ dx}{\int_0^1 x^8\ dx}\ x^4 = -\frac{9}{8} x^4 + x^3$$

$$u_3 = x_3 - \frac{u_1 \cdot x_3}{u_1 \cdot u_1} u_1 - \frac{u_2 \cdot x_3}{u_2 \cdot u_2} u_2 = x^2 - \frac{9}{7} x^4 - \frac{448}{168}\left(-\frac{9}{8} x^4 + x^3\right)$$

$$= \frac{12}{7} x^4 - \frac{8}{3} x^3 + x^2$$

I can't stand going any further. So far, my orthogonal basis is

x^4, $-\frac{9}{8} x^4 + x^3$, $\frac{12}{7} x^4 - \frac{8}{3} x^3 + x^2$, plus two more.

SOLUTIONS Section 10.3

1. (a) $\|f\|^2 = \int_1^2 |f(x)|^2\, dx = \int_1^2 (x^4 + x^2)\, dx = \frac{128}{15}$.

$\|f\| = \sqrt{\frac{128}{15}}$, normalized f is $\sqrt{\frac{15}{128}}\, x^2 + i\sqrt{\frac{15}{128}}\, x$.

(b) $\|f\|^2 = \int_0^1 |f(x)|^2\, dx = \int_0^1 (x^4 + x^2)\, dx = \frac{8}{15}$.

$\|f\| = \sqrt{\frac{15}{8}}$, normalized f is $\sqrt{\frac{15}{8}}\, x^2 + i\sqrt{\frac{15}{8}}\, x$

2. $\|1\|^2 = \int_0^L 1^2\, dx = L$ so $\|1\| = \sqrt{L}$ and the normalized function is $\frac{1}{\sqrt{L}}$.

$\|\cos \frac{7\pi x}{L}\|^2 = \int_0^L \cos^2 \frac{7\pi x}{L}\, dx = \frac{1}{2} L$ so $\|\cos \frac{7\pi x}{L}\| = \sqrt{\frac{1}{2} L}$ and the normalized function is $\sqrt{\frac{2}{L}} \cos \frac{\pi x}{L}$.

3. (a) $< f, g > = 0$ so $\int_a^b f(x)\, g(x)\, dx = 0$.

$\|f(x)\| = 1$ so $\|f(x)\|^2 = 1$ so $\int_a^b f^2(x)\, dx = 1$

And $\int_a^b g^2(x)\, dx = 1$.

(b) $\int_a^b \overline{f(x)}\, g(x)\, dx = 0$

$\int_a^b |f(x)|^2\, dx = 1$ (the vertical bars mean magntiude of the complex number f(x))

$\int_a^b |g(x)|^2\, dx = 1$ (the vertical bars mean magntiude of the complex number g(x))

4. Need $\|a\| = 1$, $\sqrt{\int_0^2 a^2\, dx} = 1$, $2a^2 = 1$, $a = \pm\sqrt{\frac{1}{2}}$

Need $< a, bx + c > = 0$, $\int_0^2 a(bx + c)\, dx = 0$, $(\frac{1}{2} bx^2 + cx)\Big|_0^2 = 0$,

$2b + 2c = 0$, $b = -c$

Need $\|bx + c\| = 1$, $\sqrt{\int_0^2 (bx + c)^2\, dx} = 1$, $\left.\frac{(bx + c)^3}{3b}\right|_0^2 = 1$,

But $b = -c$ so $\frac{2b^3}{3b} = 1$, $b = \pm\sqrt{\frac{3}{2}}$, $c = \mp\sqrt{\frac{3}{2}}$

There are four sets of answers:

$a = \sqrt{\frac{1}{2}}$, $b = \sqrt{\frac{3}{2}}$, $c = -\sqrt{\frac{3}{2}}$; $a = \sqrt{\frac{1}{2}}$, $b = -\sqrt{\frac{3}{2}}$, $c = \sqrt{\frac{3}{2}}$

$a = -\sqrt{\frac{1}{2}}$, $b = \sqrt{\frac{3}{2}}$, $c = -\sqrt{\frac{3}{2}}$; $a = -\sqrt{\frac{1}{2}}$, $b = -\sqrt{\frac{3}{2}}$, $c = \sqrt{\frac{3}{2}}$

5. If $u = (u_1, u_2)$ then the associated norm is $\|u\| = \sqrt{u \bullet u} = \sqrt{2u_1^2 + 3u_2^2}$

If $u = (3,-2)$ then the associated norm of u is $\sqrt{18 + 12} = \sqrt{30}$

6. (a) Not a norm.

reason 1 If $u = (-3,-5,1)$ then $\|u\| = -7$.
But that violates the norm axiom $\|u\| \geq 0$

reason 2 Doesn't satisfy the axiom $\|k\vec{u}\| = |k|\|\vec{u}\|$
Here's a counterexample
Let $u = (2,3)$. Then $-2u = (-4,-6)$ and $\|u\| = 5$, $2\|u\| = 10$ but $\|-2u\| = -10$.

(b) Yes. It satisfies the three norm axioms.

(c) No. If $u = (3,3,3)$ then $\|u\| = 0$. Violates the axiom "$\|u\| > 0$ if $u \neq \vec{0}$".

7. *reason 1* Determinants don't always come out ≥ 0 but norms must be ≥ 0.

reason 2 $|kM|$ doesn't always equal $|k||M|$ violating another norm axiom; in fact $|kM| = k^2|M|$ by a det rule.

reason 3 The triangle inequality doesn't hold. For a counterexample let

$A = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$. Then $A + B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$,

$|A| = 0$, $|B| = 0$, $|A + B| = 2$, and $|A + B|$ is *not* $\leq |A| + |B|$.

8. (a) $x \circ x = x_1 x_1 - x_2 x_1 - x_1 x_2 + 2x_2 x_2 = x_1^2 - 2x_1 x_2 + 2x_2^2 = (x_1 - x_2)^2 + x_2^2$

So $x \circ x$ is a sum of (real) squares so it is ≥ 0, and it *is* 0 iff both x_1 and x_2 are 0, i.e., iff $x = \vec{0}$.

(b) $i \circ j = (1)(0) - (0)(0) - (1)(1) + 2(0)(1) = -1$ Not orthog.

(c) $$u \circ v = \frac{2}{\sqrt{10}}\,\frac{4}{\sqrt{10}} - \frac{3}{\sqrt{10}}\,\frac{4}{\sqrt{10}} - \frac{2}{\sqrt{10}}\,\frac{1}{\sqrt{10}} + 2\,\frac{3}{\sqrt{10}}\,\frac{1}{\sqrt{10}} = 0$$

so u and v are orthogonal w.r.t. the new dot.

$$\|u\|_o = \sqrt{u \circ u} = \sqrt{\frac{4}{10} - \frac{6}{10} - \frac{6}{10} + \frac{18}{10}} = 1$$

$$\|v\|_o = \sqrt{v \circ v} = \sqrt{\frac{16}{10} - \frac{4}{10} - \frac{4}{10} + \frac{2}{10}} = 1$$

so u and v are unit vectors w.r.t. the norm associated with the new dot.

SOLUTIONS Section 10.4

1. (a) On [0,L], $< y_5, y_8 > = \int_0^L \cos \frac{5\pi x}{L} \cos \frac{8\pi x}{L}\, dx$

Let a computer do the integral and expect it to come out to be 0. Then you can say that y_5 and y_8 are orthog on [0,L].

(b)

$$c_0 = \frac{< 1, e^x >}{< 1, 1 >} = \frac{\int_0^2 1e^x\, dx}{\int_0^2 1^2\, dx} = \frac{e^2-1}{2}$$

$$c_1 = \frac{< \cos \frac{\pi x}{2}, e^x >}{< \cos \frac{\pi x}{2}, \cos \frac{\pi x}{2} >} = \frac{\int_0^2 e^x \cos \frac{\pi x}{2}\, dx}{\int_0^2 \cos^2 \frac{\pi x}{2}\, dx}$$

2. (a) $\int_0^1 (3-4x)x^2\, dx = \int_0^1 (3x^2 - 4x^3)\, dx = 0$ so $3-4x$ and x^2 are orthog on [0,1].

(b)

$$a = \frac{< 3-4x, e^x >}{< 3-4x, 3-4x >} = \frac{\int_0^1 (3-4x)\, e^x\, dx}{\int_0^1 (3-4x)^2\, dx}$$

$$b = \frac{< x^2, e^x >}{< x^2, x^2 >} = \frac{\int_0^1 x^2 e^x\, dx}{\int_0^1 x^4\, dx}$$

(c) The choice of a and b from part (b) minimizes the norm

$$(*) \qquad \|a(3-4x) + bx^2 - e^x\|$$

In particular, the a and b from part (b) minimize

$$\int_0^1 \left[a(3-4x) + bx^2 - e^x\right]^2 dx$$

3. The sum of the five terms is the projection of x^2 into the 5-dim subspace of function space with orthogonal (on the interval [0,3]) basis

$$\sin \frac{\pi x}{3}, \ \sin \frac{2\pi x}{3}, \ \sin \frac{3\pi x}{3}, \ \sin \frac{4\pi x}{3}, \ \sin \frac{5\pi x}{3}.$$

It's the "best" approximation to x^2 of the form $A_1 \sin \frac{\pi x}{3} + \ldots + A_5 \sin \frac{5\pi x}{3}$ in the sense that this choice of A's minimizes

$$\int_0^3 \left[x^2 - (A_1 \sin \frac{\pi x}{3} + \ldots + A_5 \sin \frac{5\pi x}{3})\right]^2 dx.$$

SOLUTIONS Section 10.5

1. Must disprove one of the linearity properties. Consider say $y = x^2$. Then
$T(3y) = (3y)(3y)' = 3x^2\ (6x) = 18x^3$
But $3T(y) = 3yy' = 3x^2\ (2x) = 6x^3$
So it isn't true that $T(ky) = kT(y)$ for all vectors y and all scalars k.
So T is not linear.

2. $T(\vec{0}) = T(u - u) = T(u) - T(u)$ (by linearity) $= \vec{0}$

3 (a) If $u_1, \ldots, u_n$ are dep then one is a combination of the others, say

$$u_1 = a_2 u_2 + \ldots + a_n u_n$$

Then

$$T(u_1) = T(a_2 u_2 + \ldots + a_n u_n)$$
$$= a_2\, T(u_2) + \ldots + a_n\, T(u_n) \quad \text{by linearity}$$

So not only are $T(u_1), \ldots, T(u_n)$ dependent, but they are related to one another in the same way that $u_1, \ldots, u_n$ were related to one another

(b) Here's one counterexample. Look at the vector space R^2 and the linear operator

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

M sends the two *independent* vectors $\vec{i}$ and $\vec{j}$ to the two *dependent* vectors (1,0) and (0,0). So M doesn't preserve independence.
Here's another counterexample. Look at function space and the differential operator $L(y) = y'$. Then L sends the two independent functions $x^3 + 3$, $x^3 + 4$ to the same output, namely $3x^2$. So L doesn't preserve independence.

4. $T\left(\sin \frac{7\pi x}{4}\right) = -\left(\frac{7\pi}{4}\right)^2 \sin \frac{7\pi x}{4}$

So $\sin \frac{7\pi x}{4}$ is an eigenfunction (because it mapped to a multiple of itself) with corresponding eigenvalue $-\left(\frac{7\pi}{4}\right)^2$.

5. (a) $< Tf, g > = \int_a^b \overline{xf(x)}\ g(x)\ dx$ by definition of the standard dot

$$= \int_a^b x\ \overline{f(x)}\ g(x)\ dx$$ by the algebra rule $\overline{zw} = \bar{z}\,\bar{w}$ and the fact that x is real so $\bar{x} = x$

On other hand,

$< f, Tg > = \int_a^b \overline{f(x)}\ xg(x)\ dx$ by definition of the standard dot

So $< Tf, g > = < f, Tg >$ which makes T self-adjoint on [a,b]

(b) $< Sf,\ g > = \int_a^b \overline{if(x)}\ g(x)\,dx = -i \int_a^b \overline{f(x)}\ g(x)\ dx$

$< f,\ Sg > = \int_a^b \overline{f(x)}\ \ ig(x)\ dx = i \int_a^b \overline{f(x)}\ g(x)\ dx$

So < Sf, g > is not always equal to < f, Sg >.
Almost any functions and any interval will serve as a counterexample (the argument isn't clinched until you produce a specific counterexample). For instance let $f(x) = x^2$, $g(x) = x^3$ and use the interval [0,1]. Then

$$< Sf,\ g > = -i \int_0^1 x^5\ dx = -\frac{1}{6}\ i$$

$$< f,\ Sg > = i \int_0^1 x^5\ dx = \frac{1}{6}\ i$$

Not equal. So it is not true that S is self-adjoint on every interval [a,b] (actually S is not self-adjoint on any interval).

6. (a) $< v, T(u) > = \overline{< T(u),\ v >}$ (dot rule)
$= 8 + i$

(b) $< u,\ T(v) > = < T(u),\ v >$ (T is self-adjoint)
$= 8 - i$

(c) Can't do with the given information.

(d) $< T(iu),\ v > = < iT(u),\ v >$ (T is linear)
$= \overline{i} < T(u),\ v >$ (dot rule)
$= -i(8-i)$
$= -1 - 8i$

(e) $< T(u+v),\ v > = < T(u) + T(v),\ v >$
$= < T(u),\ v > + < T(v),\ v >$ by the dot rule $(p+q)\cdot w = p\cdot w + q\cdot w$
$= < T(u),\ v > + < v,\ T(v) >$ since T is self-adjoint
$= 8-i + 7$
$= 15-i$

(f) $< T([2+3i]u), v> = <(2+3i)T(u),\ v>$ (T is linear)
$= (2-3i)\ <T(u), v>$ (dot rule)
$= (2-3i)(8-i)$
$= 13 - 26i$

(g) $< v,\ T([2+3i]v) > = < v,\ (2+3i)\ T(v) >$ (T is linear)
$= (2+3i) < v,\ T(v) >$ (dot rule)
$= 7(2+3i)$
$= 14 + 21i$

(h) $< T(u+v),\ v > = < T(u) + T(v),\ v >$ (T is linear)
$= [T(u) + T(v)]\cdot v$ (I'm more comfortable in the old notation)
$= T(u)\cdot v + T(v)\cdot v$ (dot rule)
$= < T(u),\ v > + < T(v),\ v >$ (back to the original notation)
$= < T(u),\ v > + < v,\ T(v) >$ (T is self-adjoint)
$= 8-i + 7$
$= 15-i$

7. *version 1* On the one hand, $< u,\ T(u) > = < T(u),\ u >$ because T self-adjoint.

On the other hand, $< u,\ T(u) > = \overline{< T(u),\ v >}$

So the number $< u,\ T(u) >$ ended up equaling a number z and also equaling $\overline{z}$.
The only way that can happen is if the number $< u,\ T(u) >$ is real.

version 2

$< u,\ T(u) > = < T(u),\ u >$ since T is self-adjoint

$= \overline{< u,\ T(u) >}$ by a dot rule

So $\langle u, T(u) \rangle$ equals its conjugate. But if a number equals its conjugate then the number is real. So $\langle u, T(u) \rangle$ is real.

$$\langle p, T(q) \rangle = \int_0^1 p(x)q'(x)\, dx$$

The two integrals are not usually equal. For instance if $p(x) = 1$ and $q(x) = x$ then $\langle T(p), q \rangle = 0$ and $\langle p, T(q) \rangle = 1$.
So T is not self-adjoint on [0,1].

(b) Need a nonzero poly p and scalar λ so that $T(p) = \lambda p$, $p' = \lambda p$. The only time that the derivative of a poly is a multiple of the poly is if the poly is a constant and $\lambda = 0$. So the eigenvectors are the polys of degree 0 (e.g., 1, 2, π, -4, 1/2, ...), all corresponding to eigenvalue 0. It's a 1-dim eigenspace.

9.(a)

$L(A + B)$	$= (A + B)^T$	definition of L
	$= A^T + B^T$	T rule
	$= L(A) + L(B)$	definition of L

$L(kA)$	$= (kA)^T$	definition of L
	$= kA^T$	T rule
	$= kL(A)$	definition of L

So L is linear.
(b) Need to find nonzero matrices A and scalars λ so that $L(A) = \lambda A$, i.e., so that $A^T = \lambda A$.

If A is symmetric then $A^T = A$. In that case, $L(A) = 1A$, making $\lambda=1$ an eigenvalue of L with all symmetric matrices as the corresponding eigenvectors.

The space of symmetric matrices has basis

$$B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

For instance, if

$$A = \begin{bmatrix} 3 & 5 \\ 5 & 7 \end{bmatrix}$$

then

$$A = 3B_1 + 7B_2 + 5B_3$$

So the eigenspace corresponding to $\lambda=1$ is 3-dim.

If A is skew-symmetric then $A^T = -A$. This makes -1 an eigenvalue of L. The corresponding eigenvectors are the skew-symmetric matrices.

The corresponding eigenspace is 1-dim with basis vector $K = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ because every 2×2 skew-symmetric matrix is a multiple of K.

(c) Let

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \text{ and } B = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$$

Then

$$L(A) = \begin{bmatrix} a_1 & a_3 \\ a_2 & a_4 \end{bmatrix}, \quad L(B) = \begin{bmatrix} b_1 & b_3 \\ b_2 & b_4 \end{bmatrix}$$

$L(A) \cdot B = a_1 b_1 + a_3 b_2 + a_2 b_3 + a_4 b_4$

$A \cdot L(B) = a_1 b_1 + a_2 b_3 + a_3 b_2 + a_4 b_4$

These are equal so L is self-adjoint.

10. Let f(x) and g(x) be any two functions. I'll find < T(f), g > and < f, T(g) > and show that they are the same.

$$< T(f),\ g > = \int_a^b T(f)\ g(x)\ dx = \int_a^b f(x)\ \sin x\ \ g(x)\ dx$$

$$< f,\ T(g) > = \int_a^b = f(x)\ T(g)\ dx = \int_a^b f(x)\ g(x)\ \sin x\ dx$$

The two integrals are the same. So T is self-adjoint on [a,b].

11. (a) Let C and D be arbitrary matrices and let α be a scalar. Then

$T(C + D) = A(C + D)$ definition of T

$= AC + AD$ matrix algebra

$= T(C) + T(D)$ definition of T again.

And

$T(\alpha C) = A(\alpha C)$ definition of T

$= \alpha(AC)$ matrix algebra

$= \alpha T(C)$ definition of T again.

So T is linear.

(b) Let $C = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$ and $D = \begin{bmatrix} d_1 & d_2 \\ d_3 & d_4 \end{bmatrix}$

I'll find $T(C) \cdot D$ and $C \cdot T(D)$ to see if they are the same.

$$T(C) = AC = \begin{bmatrix} 3c_1 & 3c_2 \\ 4c_3 & 4c_4 \end{bmatrix}$$

$$T(D) = AD = \begin{bmatrix} 3d_1 & 3d_2 \\ 4d_3 & 4d_4 \end{bmatrix}$$

$T(C) \cdot D = AC \cdot D = 3c_1\ d_1 + 3c_2\ d_2 + 4c_3\ d_3 + 4c_4\ d_4$

$C \cdot T(D) = C \cdot AD = c_1\ 3d_1 + c_2\ 3d_2 + c_3\ 4d_3 + c_4\ 4d_4$

$T(C) \cdot D$ does equal $C \cdot T(D)$ so T is self-adjoint w.r.t. this dot.

12. Find $T(x^n)$ and see if it comes out to be a multiple of x^n.

$T(x^n) = x^2\ n(n-1)x^{n-2} + x\ nx^{n-1}$

$= n(n-1)x^n + nx^n$

$= [n(n-1) + n]\ x^n$

$= n^2\ x^n$ (a multiple of x^n; the multiple is n^2)

So the corresponding eigenvalue is n^2.

SOLUTIONS Section 10.6

1. Example 1 says that the functions $\sin \frac{18\pi x}{L}$ and $\sin \frac{3\pi x}{L}$ are orthogonal on [0,L] so the integral has to be 0.

2. It's the projection of f(x) into the 10-dim subspace of function space with basis $\sin \frac{\pi x}{L}$, $\sin \frac{2\pi x}{L}$, ..., $\sin \frac{10\pi x}{L}$.

Equivalently, using the formulas in (6) for $C_1, \ldots, C_{10}$ minimizes the norm

$$\left\| f(x) - \left(C_1 \sin \frac{\pi x}{L} + C_2 \sin \frac{2\pi x}{L} + C_3 \sin \frac{3\pi x}{L} + \ldots + C_{10} \sin \frac{10\pi x}{L} \right) \right\|$$

3. $$C_0 = \frac{\langle y_0, e^x \rangle}{\langle y_0, y_0 \rangle} = \frac{\int_0^2 1 \cdot e^x \, dx}{\int_0^2 1^2 \, dx}$$

Numerator is $e^2 - 1$. Denominator is 2. $C_0 = \frac{1}{2}(e^2 - 1)$

4. (a) $$\int_0^{2\pi} \sin mx \cos mx \, dx = \left. \frac{\sin^2 mx}{2m} \right|_0^{2\pi} = \frac{\sin^2 2m\pi}{2m} - 0$$

Since m is an integer, $\sin 2m\pi = 0$.
So the integral is 0 and sin mx and cos mx are orthog on $[0,2\pi]$.

(b) $$\|1\|^2 = \int_0^{2\pi} 1^2 \, dx = 2\pi \text{ so } \|1\| = \sqrt{2\pi}$$

$$\|\cos nx\|^2 = \int_0^{2\pi} \cos^2 nx \, dx = \left. \left(\frac{x}{2} + \frac{1}{4n} \sin 2nx \right) \right|_0^{2\pi} = \pi$$

So $\|\cos nx\| = \sqrt{\pi}$.

Similarly $\|\sin nx\| = \sqrt{\pi}$.
So the normalized eigenfunction are

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos nx}{\sqrt{\pi}}, \quad \frac{\sin nx}{\sqrt{\pi}} \qquad \text{for } n = 1,2,3,\ldots$$

(c)(i) All it takes is a little rearrangement of the series in (*): For $0 \le x \le 2\pi$,

$$(**) \quad f(x) = \boxed{A_0 \sqrt{2\pi}} \frac{1}{\sqrt{2\pi}} + \boxed{A_1 \sqrt{\pi}} \frac{\cos x}{\sqrt{\pi}} + \boxed{A_2 \sqrt{\pi}} \frac{\cos 2x}{\sqrt{\pi}} + \boxed{A_3 \sqrt{\pi}} \frac{\cos 3x}{\sqrt{\pi}}$$

$$+ \ldots + \boxed{B_1 \sqrt{\pi}} \frac{\sin x}{\sqrt{\pi}} + \boxed{B_2 \sqrt{\pi}} \frac{\sin 2x}{\sqrt{\pi}} + \boxed{B_3 \sqrt{\pi}} \frac{\sin 3x}{\sqrt{\pi}} + \ldots$$

This expresses f in terms of the normalized eigenfunctions. The coeffs are boxed.

(ii) The integral is $\|f\|^2$ where the norm is taken in the space of functions on $[0,2\pi]$. Part (i) expressed f in terms of an orthonormal basis for that space. To find $\|f\|^2$, use the rule for finding norms using coords w.r.t.an orthonormal basis (see (5') in Section 10.1):

$$\int_0^{2\pi} f^2(x)\,dx = \|f\|^2 = \text{sum of squares of all the boxed coeffs in } (**)$$

$$= 2\pi A_0^2 + \pi A_1^2 + \pi A_2^2 + \pi A_3^2 + \ldots + \pi B_1^2 + \pi B_2^2 + \pi B_3^2$$