EXAMPLES OF MATRICES

1. Data Storage

$$
\begin{array}{c}
\text{Employee Number} \\
1 \\ 2 \\ 3 \\ 4 \\ 5
\end{array}
\qquad
\begin{array}{ccc}
\text{Days Absent} & \text{Attitude} & \text{Experience}
\end{array}
$$

$$
\begin{array}{c|ccc}
\text{Employee Number} & \text{Days Absent} & \text{Attitude} & \text{Experience} \\
\hline
1 & & & \\
2 & & & \\
3 & & & \\
4 & & & \\
5 & & &
\end{array}
\;
\begin{bmatrix}
1 & 1 & 1 \\
0 & 2 & 1 \\
4 & 3 & 2 \\
6 & 7 & 4 \\
9 & 10 & 8
\end{bmatrix}
$$

2. Dominance [S.S.Ulmer, "Leadership in the Michigan Supreme Court," *Judicial Decision Making,* 1963]

$$
\begin{array}{l}
\\
\text{Volker} \\ \text{Kavanaugh} \\ \text{Detmers} \\ \text{Carr} \\ \text{Smith} \\ \text{Edwards} \\ \text{Kelly} \\ \text{Black}
\end{array}
\quad
\begin{array}{c}
\begin{array}{cccccccc} V & Ka & D & C & S & E & Ke & B \end{array} \\
\mathbf{D} = \begin{bmatrix}
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
\qquad
d_{ij} = \begin{cases} 1 & \text{if } i \text{ dominates } j \\ 0 & \text{if } i \text{ does not dominate } j \end{cases}
$$

3. Markov Process [Glass & Hall, "Social Mobility in Great Britain: A Study of Intergeneration Changes in Status," *Social Mobility in Great Britain*, 1954]

$$
\begin{array}{l}
\\
\text{Upper} \\ \text{Middle} \\ \text{Lower}
\end{array}
\quad
\begin{array}{c}
\begin{array}{ccc} U & M & L \end{array} \\
\begin{bmatrix}
.448 & .484 & .068 \\
.054 & .699 & .247 \\
.011 & .503 & .486
\end{bmatrix}
\end{array}
= \mathbf{P}
$$

where p_{ij} is the probability that a son of a person working in an occupation of class i gets a job in an occupation of class j

4. System of Linear Equations

Data:

x	y	z
0	575	50
50	857	406
90	397	994
100	348	1016

Linear Model: $z = a+bx+cy$

Simultaneous Equations

$$
\begin{aligned}
a + + 575c &= 50 \\
a + 50b + 857c &= 406 \\
a + 90b + 397c &= 994 \\
a + 100b + 48c &= 1016
\end{aligned}
$$

Matrix Representation

$$
\left[\begin{array}{ccc|c}
1 & 0 & 575 & 50 \\
1 & 50 & 857 & 406 \\
1 & 90 & 397 & 994 \\
1 & 100 & 348 & 1016
\end{array}\right]
$$

SOME MATRIX ALGEBRA PROPERTIES (THEOREMS)

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$, if defined

2. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$, if defined

3. $\mathbf{A} + \mathbf{0} = \mathbf{A}$, where $\mathbf{0}$ is a zero matrix

4. $\mathbf{A} - \mathbf{A} = \mathbf{0}$, where $\mathbf{0}$ is a zero matrix

5. If $\mathbf{A} + \mathbf{B} = \mathbf{A} + \mathbf{C}$ then $\mathbf{B} = \mathbf{C}$

6. $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$

7. $k\,\mathbf{0} = \mathbf{0}$, where $\mathbf{0}$ is a zero matrix and k is any real number

8. $0\,\mathbf{A} = \mathbf{0}$, where 0 is the scalar zero and $\mathbf{0}$ is a zero matrix

9. $-(-\mathbf{A}) = \mathbf{A}$

10. $k\,(\mathbf{A} + \mathbf{B}) = k\,\mathbf{A} + k\,\mathbf{B}$

11. $k_1(k_2\,\mathbf{A}) = (k_1 k_2)\,\mathbf{A}$

12. $(k_1 + k_2)\,\mathbf{A} = k_1\,\mathbf{A} + k_2\mathbf{A}$

13. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$, if defined

14. $k\,(\mathbf{AB}) = (k\,\mathbf{A})\mathbf{B} = \mathbf{A}(k\,\mathbf{B})$

15. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$; $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$, if defined

16. $(\mathbf{A} + \mathbf{B})^2 = (\mathbf{A} + \mathbf{B})(\mathbf{A} + \mathbf{B}) = \mathbf{A}^2 + \mathbf{BA} + \mathbf{AB} + \mathbf{B}^2$

17. $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$, where $\mathbf{I}$ is the appropriate identity matrix

18. $\mathbf{0A} = \mathbf{0}$, where $\mathbf{0}$ is the appropriate zero matrix

19. $(\mathbf{A}')' = \mathbf{A}$

20. $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$

21. $(k\,\mathbf{A})' = k\mathbf{A}'$

22. $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$

<u>Cautions</u>:

1. In general, $\mathbf{AB} \neq \mathbf{BA}$

2. If $\mathbf{AB} = \mathbf{AC}$ then it is not always true that $\mathbf{B} = \mathbf{C}$

3. If $\mathbf{AB} = \mathbf{0}$ then it is not necessarily true that $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$

Example [S.S. Ulmer, "Leadership in the Michigan Supreme Court," *Judicial Decision-Making*, G. Schubert (ed.), 1963, 13-28.]

$$
\begin{array}{c|c}
 & \begin{array}{cccccccc} V & Ka & D & C & S & E & Ke & B \end{array} \\
\begin{array}{c} V \\ Ka \\ D \\ C \\ S \\ E \\ Ke \\ B \end{array} &
\begin{bmatrix}
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
= \mathbf{D} \text{ (dominance)}
$$

$$
d_{ij} = \begin{cases} 1 \text{ if i dominates j} \\ 0 \text{ otherwise} \end{cases}
$$

Justices in the Michigan Supreme Court, 1958-1960:

V = Voelker, Ka = Kavanaugh, D = Dethmers, C = Carr, S = Smith, E = Edwards, Ke = Kelly, B = Black.

Then

$$
\mathbf{D}^2 = \begin{bmatrix}
0 & 0 & 2 & 1 & 2 & 3 & 4 & 5 \\
1 & 0 & 1 & 0 & 2 & 3 & 4 & 5 \\
0 & 1 & 0 & 1 & 1 & 2 & 3 & 4 \\
1 & 0 & 0 & 0 & 1 & 2 & 3 & 4 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

so

$$
\mathbf{D} + 1/2\,\mathbf{D}^2 = \begin{bmatrix}
0 & 1 & 1 & 1.5 & 2 & 2.5 & 3 & 3.5 \\
.5 & 0 & 1.5 & 1 & 2 & 2.5 & 3 & 3.5 \\
1 & .5 & 0 & .5 & 1.5 & 2 & 2.5 & 3 \\
.5 & 0 & 1 & 0 & 1.5 & 2 & 2.5 & 3 \\
0 & 0 & 0 & 0 & 0 & 1 & 1.5 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1.5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

Simultaneous Linear Equations

We will study m linear equations in n unknowns $x_1, x_2, \cdots, x_n$.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2. \\ \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Note: we don't require $m = n$

We can write this system of equations in matrix form as **A X = B**, where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ a_{21} & a_{21} \cdots & a_{2n} \\ \\ a_{m1} & a_{m2} \cdots & a_{mn} \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \\ b_m \end{bmatrix}$$

We solve the system by considering the augmented matrix **[A|B]** and using Gaussian elimination.

<u>Examples</u>

1. $\begin{cases} 3x_1 - 5x_2 = -23 \\ x_1 + x_2 = 3 \\ 6x_1 + 2x_2 = 2 \end{cases}$ This system can be considered as giving the point of intersection of three lines in a plane. The system has a <u>unique solution</u> namely, $x_1 = -1, x_2 = 4$.

2. Given the data

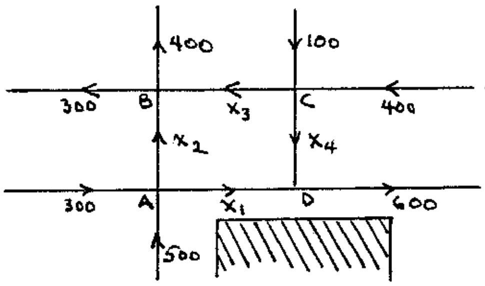
x	y	z
1	1	3
2	2	2
3	3	6

and the supposition that there is a linear relationship between x, y, and z such as $z = a + bx + cy$ for some numbers a, b, and c.

Substituting the values of x, y, z in the linear equation we obtain $\begin{cases} a + b + c = 3 \\ a + 2b + 2c = 2 \\ a + 3b + 3c = 6 \end{cases}$

This system has <u>no solution</u> (the system is <u>overdetermined</u>).

3. The figure below shows a portion of one-way streets in downtown Baltimore. The arrows indicate the direction of traffic flow and the numbers indicate the vehicle count per hour on that part of the street during a typical early afternoon. Traffic signals are located at points A, B, C, D to regulate the flow along the parts of the streets marked by x_1, x_2, x_3, x_4. We assume that the number of vehicles entering any intersection is equal to the number of vehicles leaving that intersection.



A: $300 + 500 = x_1 + x_2$

B: $x_2 + x_3 = 300 + 400$

C: $100 + 400 = x_3 + x_4$

D: $x_1 + x_4 = 600$

$$\begin{cases} x_1 + x_2 = 800 \\ x_2 + x_3 = 700 \\ x_3 + x_4 = 500 \\ x_1 + x_4 = 600 \end{cases}$$

This system has <u>many solutions</u> (the system in <u>underdetermined</u>).

Gaussian Elimination Algorithm for Finding $\mathbf{A}^{-1}$

We perform elementary row operations of three types on the augmented matrix $[\mathbf{A}|\mathbf{I}]$:

(1) Interchange two rows $(R_i \leftrightarrow R_j)$

(2) Multiply any row by a nonzero scalar $(k\,R_i)$

(3) Add a multiple of one row to another row $(k\,R_i + R_j)$

<u>Algorithm</u>

1. Start at the top row.
2. Is there a nonzero entry in the diagonal position?
   No -- if possible, interchange with a lower row to get a nonzero entry
   if not possible, then $\mathbf{A}^{-1}$ does not exist.
   Yes -- divide each element in the row by the diagonal element (*pivot*) to get a 1 in the pivot position. Note: this is the same as multiplying by 1/diagonal.
3. Add multiples of the row being considered to the lower rows to "sweep out" any nonzero entries below the pivot.
4. Go to the next lower row and repeat steps 2 and 3.
5. Continue as in 4 until all rows are considered. You will now have an *echelon matrix* (see p. 32).
6. Now begin from the bottom row and work up to sweep out nonzero entries above the pivots.
7. When you reach the form $[\mathbf{I}|\,\mathbf{B}]$ then $\mathbf{B} = \mathbf{A}^{-1}$.

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

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 2 & -2 & 0 & 0 & 1 & 0 \\ 4 & 4 & -4 & 0 & 0 & 1 \end{array}\right] \xrightarrow[-4R_1+R_3]{-2R_1+R_2} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -2 & 1 & 0 \\ 0 & 8 & -8 & -4 & 0 & 1 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 8 & -8 & -4 & 0 & 1 \\ 0 & 0 & -2 & -2 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[-1/2R_3]{1/8R_2} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1/2 & 0 & 1/8 \\ 0 & 0 & 1 & 1 & -1/2 & 0 \end{array}\right] \xrightarrow[-1R_3+R_1]{1R_3+R_2} \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/8 \\ 0 & 0 & 1 & 1 & -1/2 & 0 \end{array}\right] \xrightarrow{1R_2+R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & 0 & 1/8 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/8 \\ 0 & 0 & 1 & 1 & -1/2 & 0 \end{array}\right]$$

$$\mathbf{A}^{-1} = \begin{bmatrix} 1/2 & 0 & 1/8 \\ 1/2 & -1/2 & 1/8 \\ 1 & -1/2 & 0 \end{bmatrix} = 1/8 \begin{bmatrix} 4 & 0 & 1 \\ 4 & -4 & 1 \\ 8 & -4 & 0 \end{bmatrix}$$

Variance-Covariance Matrix

Recall: the variance of a variable X is given by $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

$$= \frac{1}{n-1} \left[\Sigma (x_i^2) - \frac{1}{n} (\Sigma x_i)^2 \right],$$ where $\bar{x}$ is the mean and the x_i are n values of X.

If we have m variables instead of one and we have n values for each of these m variables, we can record the data as an nxm matrix.

Example: $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 4 & 3 & 2 \\ 6 & 7 & 4 \\ 9 & 10 & 8 \end{bmatrix}$ where X_1 = days absent, X_2 = attitude score, and X_3 = experience. There are 5 employees. Here m = 3 (three variables) and n = 5 (five values each).

Then $\mathbf{A'A} = \begin{bmatrix} 134 & 145 & 105 \\ 145 & 163 & 117 \\ 105 & 117 & 86 \end{bmatrix}$ Note that this matrix is symmetric (see problem 15). The diagonal elements are sums of squares and the off-diagonal elements are sums of cross-products. e.g. $86 = \Sigma X_3^2$, $117 = \Sigma X_2 X_3$.

Let **U** be an nxn matrix with each $u_{ij} = 1$ (in the example n = 5). Now calculate **A′U A**. For the example

$$\mathbf{A'U\,A} = \begin{bmatrix} 1 & 0 & 4 & 6 & 9 \\ 1 & 2 & 3 & 7 & 10 \\ 1 & 1 & 2 & 4 & 8 \end{bmatrix} \begin{bmatrix} 20 & 23 & 16 \\ 20 & 23 & 16 \\ 20 & 23 & 16 \\ 20 & 23 & 16 \\ 20 & 23 & 16 \end{bmatrix}$$

Note that each column of the second matrix has the sum of the values of the corresponding variable written n times. e.g. $16 = \Sigma X_3$.

$$= \begin{bmatrix} 400 & 460 & 320 \\ 460 & 529 & 368 \\ 320 & 368 & 256 \end{bmatrix}$$

Note that this matrix is symmetric. Also $460 = (\Sigma X_1)(\Sigma X_2)$, $256 = (\Sigma X_3)^2$, etc.

The <u>mean-corrected matrix of sums of squares and cross-products</u> is given by $\mathbf{S} = \mathbf{A'A} - (1/n)\,\mathbf{A'U\,A}$.

In the example, $$\mathbf{S} = \begin{bmatrix} 134 & 145 & 105 \\ 145 & 163 & 117 \\ 105 & 117 & 86 \end{bmatrix} - (1/5) \begin{bmatrix} 400 & 460 & 320 \\ 460 & 529 & 368 \\ 320 & 368 & 256 \end{bmatrix} = \begin{bmatrix} 54 & 53 & 41 \\ 53 & 57.2 & 43.4 \\ 41 & 43.4 & 34.8 \end{bmatrix}$$

Note that $34.8 = 86 - (1/5)(16)^2 = \Sigma X_3^2 - (1/5)(\Sigma X_3)^2$

The <u>variance-covariance matrix</u> is given by $\mathbf{C} = \frac{1}{n-1} \mathbf{S}$

In the example, $$\mathbf{C} = (1/4)\,\mathbf{S} = \begin{bmatrix} 13.5 & 13.25 & 10.25 \\ 13.25 & 14.3 & 10.85 \\ 10.25 & 10.85 & 8.7 \end{bmatrix}$$

Note that $8.7 = (1/4)[86 - (1/5)(16)^2]$

$= (1/4)[\Sigma X_3^2 - (1/5)(\Sigma X_3)^2]$,

which is the variance of X_3.

Gaussian Elimination Algorithm for Finding A^{-1}

We perform elementary row operations of three types on the augmented matrix $[\mathbf{A}|\mathbf{I}]$:

(1) Interchange two rows $(R_i \leftrightarrow R_j)$

(2) Multiply any row by a nonzero scalar $(k\,R_i)$

(3) Add a multiple of one row to another row $(k\,R_i + R_j)$

<u>Algorithm</u>

1. Start at the top row.
2. Is there a nonzero entry in the diagonal position?
   No -- if possible, interchange with a lower row to get a nonzero entry
   if not possible, then $\mathbf{A}^{-1}$ does not exist.
   Yes -- divide each element in the row by the diagonal element (*pivot*) to get a 1 in the pivot position. Note: this is the same as multiplying by 1/diagonal.
3. Add multiples of the row being considered to the lower rows to "sweep out" any nonzero entries below the pivot.
4. Go to the next lower row and repeat steps 2 and 3.
5. Continue as in 4 until all rows are considered. You will now have an *echelon matrix* (see p. 32).
6. Now begin from the bottom row and work up to sweep out nonzero entries above the pivots.
7. When you reach the form $[\mathbf{I}|\,\mathbf{B}]$ then $\mathbf{B} = \mathbf{A}^{-1}$.

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

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 2 & -2 & 0 & 0 & 1 & 0 \\ 4 & 4 & -4 & 0 & 0 & 1 \end{array}\right] \xrightarrow[-4R_1+R_3]{-2R_1+R_2} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -2 & 1 & 0 \\ 0 & 8 & -8 & -4 & 0 & 1 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 8 & -8 & -4 & 0 & 1 \\ 0 & 0 & -2 & -2 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[-1/2R_3]{1/8R_2} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1/2 & 0 & 1/8 \\ 0 & 0 & 1 & 1 & -1/2 & 0 \end{array}\right] \xrightarrow[-1R_3+R_1]{1R_3+R_2} \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/8 \\ 0 & 0 & 1 & 1 & -1/2 & 0 \end{array}\right] \xrightarrow{1R_2+R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & 0 & 1/8 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/8 \\ 0 & 0 & 1 & 1 & -1/2 & 0 \end{array}\right]$$

$$\mathbf{A}^{-1} = \begin{bmatrix} 1/2 & 0 & 1/8 \\ 1/2 & -1/2 & 1/8 \\ 1 & -1/2 & 0 \end{bmatrix} = 1/8 \begin{bmatrix} 4 & 0 & 1 \\ 4 & -4 & 1 \\ 8 & -4 & 0 \end{bmatrix}$$

The Normal Equations

Given m independent variables $X_1, X_2, \cdots, X_m$ and the dependent variable Y. We want to find a linear relationship (regression equation) $Y = b_0 + b_1X_1 + b_2X_2 + \cdots + b_mX_m$. Suppose that for each of the variables we have n values (observations). We can then write ther system of equations $\mathbf{X}\,\mathbf{B} = \mathbf{Y}$, where

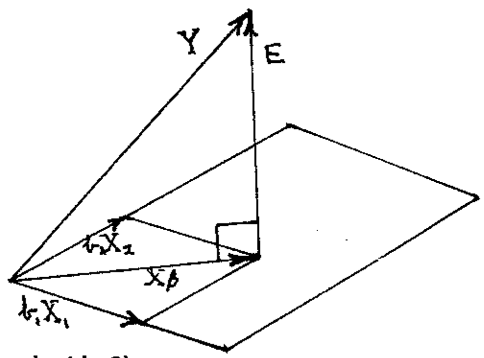
$$\mathbf{X} = \underset{n\mathrm{x}(m+1)}{\begin{bmatrix} 1 & | & | & & | \\ 1 & | & | & & | \\ \vdots & X_1 & X_2 & \cdots & X_m \\ \cdot & | & | & & | \\ 1 & | & | & & | \end{bmatrix}} \qquad \mathbf{Y} = \underset{n\mathrm{x}1}{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}} \qquad \mathbf{B} = \underset{(m+1)\mathrm{x}1}{\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}}$$

Typically the system has no solution (it is overdetermined) so we try to find a matrix $\hat{\mathbf{B}}$ so that $\mathbf{Y} = \mathbf{X}\,\hat{\mathbf{B}} + \mathbf{E}$, where

$$\mathbf{E} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

is the error term and $\hat{\mathbf{B}}$ is the best approximation to a solution in the sense that the length of the vector $\mathbf{E}$ is minimized, that is, $\sqrt{\mathbf{E}'\cdot\mathbf{E}} = \sqrt{\Sigma e_i^2}$ is as small as possible. This is the idea of "least squares."

We will look at a geometric argument to obtain $\hat{\mathbf{B}}$. Recall that the shortest distance from a point to a plane is measured along a perpendicular line from the point to the plane. Also recall that a line will be perpendicular to a plane if it is perpendicular (orthogonal) to every line in the plane passing through the point of intersection of the line and the plane. Using the generalization of these geometric facts, we want to choose $\hat{\mathbf{B}}$ so that the vector $\mathbf{E}$ is orthogonal to the vector $\mathbf{X}\mathbf{B}$ for every choice of $\mathbf{B}$. That is, we want to have



$(\mathbf{X}\,\mathbf{B})'\cdot\mathbf{E} = 0$	(vectors are orthogonal when the inner product is 0)
$(\mathbf{X}\,\mathbf{B})'\cdot(\mathbf{Y} - \mathbf{X}\,\hat{\mathbf{B}}) = 0$	(since from above we want $\mathbf{Y} = \mathbf{X}\,\hat{\mathbf{B}} + \mathbf{E}$)
$(\mathbf{X}\,\mathbf{B})'(\mathbf{Y} - \mathbf{X}\,\hat{\mathbf{B}}) = 0$	(writing the inner product as matrix multiplication)
$(\mathbf{B}'\,\mathbf{X}')(\mathbf{Y} - \mathbf{X}\,\hat{\mathbf{B}}) = 0$	(a property of the transpose of a matrix)
$\mathbf{B}'[\mathbf{X}'(\mathbf{Y} - \mathbf{X}\,\hat{\mathbf{B}})] = 0$	(a property of matrix multiplication)
$\mathbf{B}'[\mathbf{X}'\mathbf{Y} - \mathbf{X}'\mathbf{X}\,\hat{\mathbf{B}}] = 0$	(a property of matrix multiplication)

Since this last statement must be true for <u>every</u> choice of $\mathbf{B}$, we must have $\mathbf{X}'\mathbf{Y} - \mathbf{X}'\mathbf{X}\hat{\mathbf{B}} = \mathbf{0}$.
This gives us the <u>normal equations</u>

$$\boxed{\mathbf{X}'\mathbf{X}\hat{\mathbf{B}} = \mathbf{X}'\mathbf{Y}}$$

Note that if $\mathbf{X}'\mathbf{X}$ is non-singular (invertible) then $\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.
Even if it is invertible, we can avoid calculating the inverse by solving the system of equations directly, i.e.

$$[\mathbf{X}'\mathbf{X} \mid \mathbf{X}'\mathbf{Y}] \longrightarrow \cdots \longrightarrow [\mathbf{I} \mid \hat{\mathbf{B}}]$$

EXAMPLE Recall the data given earlier

X_1	X_2	Y
1	1	1
2	1	0
3	2	4
7	4	6
10	8	9

X_1 = attitude score

X_2 = years experience

Y = days absent

Suppose we think $Y = b_0 + b_1 X_1 + b_2 X_2$

$$\therefore X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 7 & 4 \\ 1 & 10 & 8 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 \\ 0 \\ 4 \\ 6 \\ 9 \end{bmatrix}, \quad \beta = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}.$$

Note that $X\beta = Y$ has no solution so we find $\hat{\beta}$.

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 7 & 10 \\ 1 & 1 & 2 & 4 & 8 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 7 & 4 \\ 1 & 10 & 8 \end{bmatrix} = \begin{bmatrix} 5 & 23 & 16 \\ 23 & 163 & 117 \\ 16 & 117 & 86 \end{bmatrix}$$

Note this matrix is symmetric.

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 7 & 10 \\ 1 & 1 & 2 & 4 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 4 \\ 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 20 \\ 145 \\ 105 \end{bmatrix}$$

To find $\hat{\beta}$ we solve $X'X\hat{\beta} = X'Y$, i.e. $\begin{bmatrix} 5 & 23 & 16 \\ 23 & 163 & 117 \\ 16 & 117 & 86 \end{bmatrix} \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} = \begin{bmatrix} 20 \\ 145 \\ 105 \end{bmatrix}$

Here $$(X'X)^{-1} = \frac{1}{535} \begin{bmatrix} 329 & -106 & 83 \\ -106 & 174 & -217 \\ 83 & -217 & 286 \end{bmatrix}$$

$$\therefore \hat{\beta} = (X'X)^{-1} X'Y = \frac{1}{535} \begin{bmatrix} 329 & -106 & 83 \\ -106 & 174 & -217 \\ 83 & -217 & 286 \end{bmatrix} \begin{bmatrix} 20 \\ 145 \\ 105 \end{bmatrix} = \frac{1}{535} \begin{bmatrix} -75 \\ 325 \\ 225 \end{bmatrix} \approx \begin{bmatrix} -0.14 \\ 0.61 \\ 0.42 \end{bmatrix}$$

$$\therefore \hat{Y} = -0.14 + 0.61 X_1 + 0.42 X_2$$

Alternately, we could solve the system $(X'X)\hat{\beta} = X'Y$ by Gaussian elimination

$$\left[\begin{array}{ccc|c} 5 & 23 & 16 & 20 \\ 23 & 163 & 117 & 145 \\ 16 & 117 & 86 & 105 \end{array}\right] \longrightarrow \cdots \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -0.140 \\ 0 & 1 & 0 & 0.607 \\ 0 & 0 & 1 & 0.421 \end{array}\right]$$

rounding off to three decimal places.

THE DETERMINANT OF A MATRIX

A must be a square matrix. The determinant of **A**, written det **A** or |**A**|, is a single real number assigned to **A**. (Recall that the trace of **A**, as defined in problem 37, is also a real number assigned to **A** in a certain way.) At one time in history when matrices were small, determinants played a major role in solving systems of linear equations, in analytic geometry, and in other parts of mathematics. Today, when we study such large matrices, determinants do not have such a central computational role. However, they do give important information about matrices and a knowledge of them is useful in some applications (for example, the determinant of a variance-covariance matrix gives the concept of a generalized variance in multivariate analysis, the determinant is used to understand eigenvalues in principal component analysis, etc.)

We now give a recursive definition of the determinant.

For a 1x1 matrix $\mathbf{A} = [a_{11}]$, det $\mathbf{A} = a_{11}$

For a 2x2 matrix $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, det $\mathbf{A} = a_{11}a_{22} - a_{12}a_{21}$

For an nxn matrix,

$$\det \mathbf{A} = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in},$$

where the cofactors $C_{ij} = (-1)^{i+j} \det \mathbf{A}_{ij}$
and $\mathbf{A}_{ij}$ is the submatrix obtained from **A** by deleting the ith row and the jth column of **A**.

[This is called the expansion of det **A** across the ith row.] Also

$$\det \mathbf{A} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

[This is called the expansion of det **A** down the jth column.]

Note that the recursive definition gives the determinant of a 3x3 matrix in terms of determinants of 2x2 matrices, the determinant of a 4x4 matrix in terms of determinants of 3x3 matrices, etc.

Example: Find the determinant of the matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 & 5 \\ 4 & 0 & 3 \\ 2 & 1 & 4 \end{bmatrix}$

(a) Expansion along the first row

$$\det \mathbf{A} = 0C_{11} + 1C_{12} + 5C_{13}$$

$$C_{12} = (-1)^{1+2} \det \mathbf{A}_{12} = (-1) \begin{vmatrix} 4 & 3 \\ 2 & 4 \end{vmatrix} = (-1)(16-6) = -10; \quad C_{13} = (-1)^{1+3} \det \mathbf{A}_{13} = (+1) \begin{vmatrix} 4 & 0 \\ 2 & 1 \end{vmatrix} = 4$$

Therefore, det $\mathbf{A} = 1(-10) + 5(4) = 10$

(b) Expansion down the second column

$$\det \mathbf{A} = 1C_{12} + 0C_{22} + 1C_{32} = 1(-1)^3 \begin{vmatrix} 4 & 3 \\ 2 & 4 \end{vmatrix} + 1(-1)^5 \begin{vmatrix} 0 & 5 \\ 4 & 3 \end{vmatrix} = -1(16-6) + (-1)(0-20) = 10$$

Properties of Determinants

1. If **A** is upper-triangular, that is,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ 0 & 0 & 0 & & a_{nn} \end{bmatrix}$$

then det **A** $= a_{11}a_{22}a_{33} \cdots a_{nn}$

2. If **B** is obtained from **A** by adding a multiple of one row of **A** to another row of **A**, then det **B** = det **A**.

3. If **B** is obtained from **A** by interchanging two rows of **A**, then det **B** = - det **A**.

4. If **B** is obtained from **A** by multiplying a row of **A** by $k \neq 0$, then det **B** = k det **A**.

Example Use the properties to find the determinant of

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 5 \\ 4 & 0 & 3 \\ 2 & 1 & 4 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 5 \\ 4 & 0 & 3 \\ 2 & 1 & 4 \end{bmatrix} \underset{\ominus}{\longrightarrow} \begin{bmatrix} 2 & 1 & 4 \\ 4 & 0 & 3 \\ 0 & 1 & 5 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 & 4 \\ 0 & -2 & -5 \\ 0 & 1 & 5 \end{bmatrix} \underset{\ominus}{\longrightarrow} \begin{bmatrix} 2 & 1 & 4 \\ 0 & 1 & 5 \\ 0 & -2 & -5 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 & 4 \\ 0 & 1 & 5 \\ 0 & 0 & 5 \end{bmatrix} = \mathbf{B}$$

det **A** = (-1)(-1) det **B** = (2)(1)(5) = 10 [This is what we found using the definition.]

Some Other Properties of Determinants

1. det **A**′ = det **A** Note: this means that properties 2,3,4 above are true if row is replaced by column.

2. If a row (column) of **A** consists of all zeros, then det **A** = 0.

3. If a row (column) of **A** is a multiple of another row (column) of **A**, then det **A** = 0.

4. If **A** is an nxn matrix and rank **A** < n, then det **A** = 0.

5. det **AB** = (det **A**)(det **B**).

6. det $\mathbf{A}^{-1}$ = 1/ det **A** if $\mathbf{A}^{-1}$ exists.

A^{-1} Using Cofactors

There is a method for finding the inverse of a matrix using cofactors. However, this method has more theoretical than practical value unless the matrix is small (e.g. 2x2 or 3x3). Since it is found in many texts it is good to be familiar with it. Here's how it works.

Given $\mathbf{A} = [a_{ij}]$

1. Form the matrix of cofactors: $\mathbf{C} = [A_{ij}]$, where A_{ij} is the cofactor of a_{ij}.
2. Form the adjoint matrix of **A**: adj $\mathbf{A} = \mathbf{C}'$
3. Find det **A**
4. If det $\mathbf{A} \neq 0$ then $\mathbf{A}^{-1} = (1/\det \mathbf{A})$ adj **A**

<u>Example</u>

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

1. $\mathbf{C} = \begin{bmatrix} 8 & 8 & 16 \\ 0 & -8 & -8 \\ 2 & 2 & 0 \end{bmatrix}$

A_{32} = cofactor of a_{32}

$= (-1)^{3+2} \det \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} = (-1)(0 - 2) = 2$

2. adj $\mathbf{A} = \mathbf{C}' = \begin{bmatrix} 8 & 0 & 2 \\ 8 & -8 & 2 \\ 16 & -8 & 0 \end{bmatrix}$

3. det $\mathbf{A} = 16$

4. $\mathbf{A}^{-1} = (1/16) \begin{bmatrix} 8 & 0 & 2 \\ 8 & -8 & 2 \\ 16 & -8 & 0 \end{bmatrix}$

<u>Note</u>: For an nxn matrix it requires the computation of n^2+1 determinants. So if **A** is a 10x10 matrix we need the calculation of 100 determinants of 9x9 matrices (step 1) and one additional determinant of a 10x10 matrix (step 3). For a 10x10 matrix there are approximately 3000 times as many arithmetic steps as is required by the Gaussian elimination method we learned earlier so we probably all prefer the Gaussian elimination method!

Cramer's Rule for Solving AX = B

Given the system of linear equations $\mathbf{AX} = \mathbf{B}$ where **A** is a square matrix (i.e. the number of equations and the number of unknowns are equal) and det $\mathbf{A} \neq 0$. Then to solve the system we find each variable

$x_i = (\det \mathbf{A}_i) / (\det \mathbf{A})$, where $\mathbf{A}_i$ is the matrix obtained from **A** by replacing the ith column of **A** by the column **B**.

<u>Example</u>:

$$x_1 - 2x_2 = 7$$
$$3x_1 + x_2 = 5$$

$$x_1 = \frac{\begin{vmatrix} 7 & -2 \\ 5 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & -2 \\ 3 & 1 \end{vmatrix}} = 17/7 \qquad x_2 = \frac{\begin{vmatrix} 1 & 7 \\ 3 & 5 \end{vmatrix}}{\begin{vmatrix} 1 & -2 \\ 3 & 1 \end{vmatrix}} = -16/7$$

<u>Note</u>: For n unknowns this method requires the computation of n + 1 determinants! For example, if there are 10 equations and 10 unknowns we would need to compute 11 determinants of 10x10 matrices. The use of Gaussian elimination is computationally shorter. Also Gaussian elimination does not require **A** to be square or for det **A** to be not zero. So we probably all prefer the Gaussian elimination method for solving the linear system $\mathbf{AX} = \mathbf{B}$.

Eigenvalues and Eigenvectors

Definitions: $\mathbf{A}$ is an nxn (square) matrix. An *eigenvalue* of $\mathbf{A}$ is a number λ such that it is possible to find a non-zero vector $\mathbf{X}$ with the property that $\mathbf{AX} = \lambda\mathbf{X}$. The vector $\mathbf{X}$ is called an *eigenvector* of $\mathbf{A}$ associated with the eigenvalue λ.

Note: If $\mathbf{X}$ is an eigenvector of $\mathbf{A}$ associated with λ then $k\mathbf{X}$ is too for any scalar k.

How to Find Eigenvalues

If $\mathbf{AX} = \lambda\mathbf{X}$ then $\mathbf{AX} - \lambda\mathbf{X} = \mathbf{0}$ so $\mathbf{AX} - \lambda\mathbf{I}\mathbf{X} = \mathbf{0}$ or $(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0}$. Since $\mathbf{X} \neq \mathbf{0}$, this homogeneous system of equations must have rank less than n and so $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. Note that this is equivalent to saying that $(\mathbf{A} - \lambda\mathbf{I})^{-1}$ does not exist or that $\mathbf{A} - \lambda\mathbf{I}$ is singular. We call $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ the *characteristic equation* for $\mathbf{A}$ and we solve this equation to find the eigenvalues λ.

Example: $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. $\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}$ so $\det(\mathbf{A} - \lambda\mathbf{I}) = (2-\lambda)^2 - 1$

If $(2-\lambda)^2 - 1 = 0$ then $\lambda^2 - 4\lambda + 3 = 0$ or $(\lambda - 3)(\lambda - 1) = 0$.
So $\lambda = 3, 1$ are the eigenvalues of $\mathbf{A}$.

How to Find the Associated Eigenvectors

We solve the homogeneous system of equations $(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0}$ for $\mathbf{X}$.

Example: For the matrix $\mathbf{A}$ above we found the eigenvalues $\lambda = 3$ and $\lambda = 1$.

For $\lambda = 3$ we have $(\mathbf{A} - 3\mathbf{I})\mathbf{X} = \mathbf{0}$ or $\left[\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$

So $\mathbf{X}_1 = \begin{bmatrix} k \\ k \end{bmatrix} = k\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

For $\lambda = 1$ we have $\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$ So $\mathbf{X}_2 = \begin{bmatrix} -k \\ k \end{bmatrix} = k\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

We could always get unit eigenvectors (see problem 8h) by taking k properly. In this example, taking $k = 1/\sqrt{2}$ we get $\mathbf{X}_1$ and $\mathbf{X}_2$ to be linearly independent, orthogonal, unit vectors.

Application: Principal Components

Suppose that $\mathbf{C} = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 5 & 4 \\ 3 & 4 & 5 \end{bmatrix}$ is the variance-covariance matrix for three variables X_1, X_2, and X_3 (recall an earlier handout).

The eigenvalues of $\mathbf{C}$ are found to be 12, 3, 1 (arranged from largest to smallest) and the corresponding unit eigenvectors are

$$\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

From $\mathbf{C}$ we see that the total variance is the sum of the variances of the three variables, namely, $6 + 5 + 5 = 16$.

Observe: • this sum is the trace of $\mathbf{C}$ (see problem 37 for the definition of trace).
• this sum is the sum of the eigenvalues ($12 + 3 + 1 = 16$).

This observation will always hold.

Now consider the following linear combination of the variables X_1, X_2, X_3:

$$Y_1 = 1/\sqrt{3}\ X_1 + 1/\sqrt{3}\ X_2 + 1/\sqrt{3}\ X_3$$

where the scalars are the entries in the first eigenvector above.

The variable Y_1 is called the first principal component. It will account for 75% of the total variance since it comes from the largest eigenvalue 12 and $12/16 = 0.75$.

The variable $Y_2 = -2/\sqrt{6}\ X_1 + 1/\sqrt{6}\ X_2 + 1/\sqrt{6}\ X_3$ is the second principal component.

The first and second principal components together account for almost 94% of the total variance since $\dfrac{12 + 3}{16} = 0.9375$.

MATHEMATICS FOR SOCIAL SCIENTISTS II

MATRIX ALGEBRA EXERCISES

1. The following matrix shows the average daily traffic (in thousands rounded to the nearest 1000) between five regional municipalities in Canada for the year 1996. (Data from the Joint Program in Transportation, University of Toronto)

$$
\begin{array}{l}
\text{From:} \\
\\
\text{1. Metro Toronto} \\
\text{2. Durham} \\
\text{3. York} \\
\text{4. Peel} \\
\text{5. Halton}
\end{array}
\quad
\begin{array}{c}
\text{To:} \\
\begin{array}{ccccc} \text{MT} & \text{D} & \text{Y} & \text{P} & \text{H} \end{array} \\
\begin{bmatrix}
4010 & 95 & 319 & 281 & 43 \\
95 & 725 & 24 & 6 & 1 \\
318 & 25 & 769 & 38 & 4 \\
284 & 6 & 39 & 1189 & 71 \\
43 & 1 & 4 & 71 & 528
\end{bmatrix}
\end{array}
$$

(a) What is the average daily traffic from Durham to Peel?

(b) What is the average daily traffic from Halton to Metro Toronto?

(c) What is the internal traffic in Peel?

(d) Between which two municipalities is there the largest daily traffic?

(e) What is the total number of vehicles that leave York for the other four municipalities?

(f) What is the total number of vehicles that enter Metro Toronto from the other four municipalities?

2. The following transition matrix records the expenditure quintiles for 4,304 households in Vietnam for 1993 and then in 1998. Quintile 1 is the poorest and quintile 5 is the richest. (Data from the World Bank)

$$
\begin{array}{c}
\text{Expenditure Quintile 1993} \\
1 \\ 2 \\ 3 \\ 4 \\ 5
\end{array}
\quad
\begin{array}{c}
\text{Expenditure Quintile 1998} \\
\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{bmatrix}
448 & 230 & 124 & 51 & 8 \\
237 & 259 & 217 & 115 & 33 \\
113 & 207 & 218 & 230 & 93 \\
47 & 125 & 214 & 282 & 193 \\
16 & 40 & 88 & 183 & 533
\end{bmatrix}
\end{array}
\quad
\begin{array}{c}
\text{total} \\
861 \\ 861 \\ 861 \\ 861 \\ \underline{860} \\ 4{,}304
\end{array}
$$

(a) What percent of these households moved from the poorest to the richest quintile?

(b) What percent of these households remained in the same quintile?

(c) Which quintile seems to have the highest chance of staying in the same quintile?

(d) What percent of the households moved either up or down one quintile?

3. Find values of a, b, and c so that $\begin{bmatrix} a-5 & b+3 \\ 6 & 2 \end{bmatrix} = \begin{bmatrix} 7 & -3 \\ 6 & c \end{bmatrix}$

4. Given the matrices $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \\ -4 & 5 \end{bmatrix}$ $\mathbf{C} = \begin{bmatrix} -2 & 1 \\ 0 & -3 \\ 4 & -5 \end{bmatrix}$

Find: (a) **A** + **B** (b) (**A** + **B**) + **C** (c) **A** + (**B** + **C**) (d) **A** - **C** (e) **B** + **0** (f) **A** - **A**

5. The matrix **P** gives the energy production in 1997 in three non-renewable sources for four countries. The matrix **C** gives the energy consumption in 1997 for these countries. The numbers are the energy equivalents of a billion metric tons of oil. The sources are: Solid fuels (e.g. coal), Liquid fuels (e.g. oil), and Gas (natural gas). (Data from the World Resources Institute)

$$\begin{array}{l} \\ \text{Australia} \\ \text{Germany} \\ \text{Iran} \\ \text{United States} \end{array} \quad \mathbf{P} = \begin{array}{ccc} \text{Solid} & \text{Liquid} & \text{Gas} \\ \end{array} \begin{bmatrix} 139.0 & 27.9 & 25.6 \\ 70.2 & 3.5 & 16.1 \\ 0.6 & 184.6 & 38.4 \\ 561.9 & 396.6 & 442.1 \end{bmatrix}$$

$$\begin{array}{l} \text{Australia} \\ \text{Germany} \\ \text{Iran} \\ \text{United States} \end{array} \quad \mathbf{C} = \begin{bmatrix} 42.3 & 35.6 & 16.9 \\ 86.3 & 139.3 & 71.9 \\ 0.9 & 67.6 & 38.3 \\ 513.3 & 854.5 & 508.0 \end{bmatrix}$$

Find the matrix **P** - **C** and interpret the entries.

6. Given the matrices $\mathbf{A} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 0 \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} 3 & -5 & 7 \\ -2 & 4 & -6 \end{bmatrix}$

Find:

(a) 3 **A** (b) **A** + 2 **B** (c) (1/2) **A** - **B**

Solve for **X**: (d) **A** + **X** = **B** (e) **B** - 2 **X** = **A**

7. For the matrices **A** and **B** as in Problem 6 and for the numbers p = 5 and q = -2, show that

(a) (p + q) **A** = p **A** + q **A** (b) p (**A** + **B**) = p **A** + p **B**

(c) (pq) **A** = p(q **A**) (d) (-1) **A** = -**A** (e) 0 **A** = **0**
[0 a scalar, **0** a matrix]

8. Given the vectors $\mathbf{a} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ $\mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$ $\mathbf{c} = \begin{bmatrix} -2 \\ x \\ 4 \end{bmatrix}$

Find:

(a) $\mathbf{a} + \mathbf{b}$ (b) $2\,\mathbf{a} + 3\,\mathbf{b}$ (c) $\mathbf{a} - \mathbf{b}$ (d) $\mathbf{a}' \cdot \mathbf{b}$ (e) $\mathbf{a}' \cdot \mathbf{a}$

Determine:

(f) x so that **a** and **c** are orthogonal vectors

(g) the length of **b**

(h) a unit vector (that is, one with length 1) having the same direction as **b** (that is, a positive scalar multiple of **b**).

9. Compute the following multiplications:

(a) $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ (d) $\begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -4 \end{bmatrix}$

(e) $\begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

10. Given the matrices $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 4 \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$ $\mathbf{C} = \begin{bmatrix} -1 & 2 \\ 3 & -4 \end{bmatrix}$

Compute (if possible):

(a) **AB** (b) **BA** (c) **CA** (d) **AC**

11. For the matrices $\mathbf{A} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -2 & 1 \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{bmatrix}$ $\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$

Find:

(a) $\mathbf{A}(\mathbf{BC})$ (b) $(\mathbf{AB})\mathbf{C}$ (c) $\mathbf{A}(\mathbf{B} + \mathbf{C})$ (d) $\mathbf{AB} + \mathbf{AC}$

(e) $\mathbf{A}^2$ (f) $\mathbf{I}\,\mathbf{B}$ (g) $\mathbf{0}\,\mathbf{C}$

12. Consider the matrices

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ with ad-bc} \neq 0 \text{ and}$$

$$\mathbf{B} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Find **AB** and **BA**. Note that for a scalar k, $\mathbf{A}(k\,\mathbf{B}) = k\,(\mathbf{AB})$.

13. Using $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 4 \end{bmatrix}$

verify that

(a) $(\mathbf{A}')' = \mathbf{A}$ (b) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$ (c) $(3\,\mathbf{B})' = 3\,\mathbf{B}'$

14. If **A** and **B** are arbitrary 2x2 matrices, which of the following statements are always true? If a statement is not always true, give an example to illustrate why it may not be true.

(a) $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + 2\,\mathbf{AB} + \mathbf{B}^2$ (b) $(\mathbf{A} - \mathbf{B})^2 = \mathbf{A}^2 + \mathbf{B}^2 - \mathbf{AB} - \mathbf{BA}$

(c) $(\mathbf{AB})^2 = \mathbf{A}^2\mathbf{B}^2$ (d) $(\mathbf{AB})' = \mathbf{A}'\mathbf{B}'$ (e) $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$

15. A matrix is called a <u>symmetric matrix</u> whenever $\mathbf{A}' = \mathbf{A}$. Find the values of a, b, and c if the following are symmetric matrices.

(a) $\begin{bmatrix} 3 & a & 5 \\ -1 & b & 2 \\ c & 2 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} -1 & 2 & a \\ b & 3 & c \\ -2 & 0 & 7 \end{bmatrix}$

16. In a certain district there are 20,000 voters who are registered as either Democrat or Republican. The percentage of each by different age groups is given by the following matrix:

$$\begin{array}{l c} & \begin{array}{cc} \text{Democrat} & \text{Republican} \end{array} \\ \begin{array}{l} \text{Under 30} \\ \text{30 - 50} \\ \text{Over 50} \end{array} & \begin{bmatrix} .70 & .30 \\ .60 & .40 \\ .40 & .60 \end{bmatrix} = \mathbf{A} \end{array}$$

The distribution of registered voters by age is given by the following matrix:

$$\begin{array}{cccc} & \text{Under 30} & \text{30-50} & \text{Over 50} \\ \mathbf{B} = & [\,6000 & 9000 & 5000\,] \end{array}$$

(a) Interpret the entries in the matrix **BA.**

(b) If an election were held, which party would be expected to win, assuming that the voters select the candidate of the party in which they are registered?

(c) What would be the expected result if $\mathbf{B} = [\,3000 \quad 5000 \quad 12000\,]$?

17. According to the California Current Population Survey, during the past decade a greater number of persons annually leave California for other states than enter California from another state. From the data for 2001 we estimate the fractions of the population staying or moving between California (CA) and the rest of the United States (US) and enter them in the matrix **M**.

$$\begin{array}{l c} \text{From\textbackslash To} & \begin{array}{cc} \text{CA} & \text{US} \end{array} \\ \begin{array}{l} \text{CA} \\ \text{US} \end{array} & \begin{bmatrix} .981 & .019 \\ .002 & .998 \end{bmatrix} = \mathbf{M} \end{array}$$

The population distribution of CA and US in March 2001 is approximated by

$$\mathbf{P} = [\,12.6 \qquad 87.4\,].$$

Assuming this pattern continues (i.e. we have a Markov process) determine

(a) The population distribution in 2002.

(b) The population distribution in 2003.

18. Write each of the following systems of equations as a single matrix equation of the form **AX** = **B**.

(a)
$$\begin{aligned} 2x - 3y + 4z &= 5 \\ 3x + 4y + 2z &= 7 \\ 4x - 2y - 3z &= 9 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\ 2x_1 + x_2 \qquad\quad - 3x_4 &= -1 \\ 3x_1 - x_2 + 2x_3 + x_4 &= 2 \end{aligned}$$

(c)
$$\begin{aligned} x + y &= 1 \\ 3x - 5y &= 0 \\ 2x + 7y &= 2 \end{aligned}$$

19. Recently Malta has shown a high turnout in parliamentary elections as the following data from the Electoral Office indicates.

Year	1951	1962	1971	1981	1992	1998
X = time since 1951	0	11	20	30	41	47
$Y = \dfrac{\text{number voting}}{\text{number voting age}}$	60.4	79.4	77.6	86.0	95.3	95.9

Suppose there is a straight-line regression relationship between X and Y. Write the matrix equation $\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \mathbf{e}$ for these observations (see Namboodiri p.23).

20. [Optional--it comes up in multivariate analysis]

(a) Compute $\mathbf{XAX}'$ where $\mathbf{X} = [x \;\; y \;\; z]$ and
$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 3 & 1 \\ 0 & 1 & -4 \end{bmatrix}$$

Note that **A** is a symmetric matrix.

(b) Write the quadratic polynomial $4x^2 + 6xy + y^2 - 8xz - z^2$ as a matrix product as in (a).
(Hint: get a symmetric matrix from the coefficients.)

21. [Optional; do if you want to see the variance-covariance matrix]

An experimental class was given a pretest and a final test. The scores on the two tests are as follows:

X_1 = pretest	24	23	17	13	26	25	10	13	10	13
X_2 = final test	52	43	43	19	35	64	19	24	42	48

Obtain the variance-covariance matrix **C** for the variables X_1 and X_2.

22. Show that for the matrix $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$

it is not possible to find a 2x2 matrix **B** so that $\mathbf{AB} = \mathbf{I}$, where $\mathbf{I}$ is the 2x2 identity matrix.

23. Find, if possible, the inverse of each of the following matrices:

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

24. Consider the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}$$

(a) $\mathbf{A}^{-1}$ is found in Exercise 23 (a). What is $\mathbf{B}^{-1}$? (This is easy; see how **B** is related to **A**.)

(b) Find $\mathbf{C}^{-1}$.

(c) Calculate $\mathbf{AC}$, $(\mathbf{AC})^{-1}$, and $\mathbf{C}^{-1}\mathbf{A}^{-1}$. Notice the relationship between two of these.

(d) Show that $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$.

25. A matrix A is called an <u>orthogonal matrix</u> if $\mathbf{A}^{-1} = \mathbf{A}'$ (i.e. if $\mathbf{AA}' = \mathbf{A}'\mathbf{A} = \mathbf{I}$). Find values of x so that the matrix

$$A = \begin{bmatrix} x & -2x \\ 2x & x \end{bmatrix}$$

is an orthogonal matrix.

26. [Optional; do if you know some trigonometry]

Find $\mathbf{A}^{-1}$ if $\mathbf{A} = \begin{bmatrix} \cos ø & -\sin ø \\ \sin ø & \cos ø \end{bmatrix}$

27. (a) If $\mathbf{AB} = \mathbf{C}$ and $\mathbf{A}$ is invertible, find $\mathbf{B}$.

(b) If $\mathbf{B} = \mathbf{PAQ}$, where $\mathbf{P}$ and $\mathbf{Q}$ are invertible, find $\mathbf{A}$.

(c) Given that $\mathbf{AX} = \mathbf{B} + 2\,\mathbf{X}$. State conditions on $\mathbf{A}$ and $\mathbf{B}$ that will enable us to solve for $\mathbf{X}$. Solve for $\mathbf{X}$ under these conditions.

(d) Repeat as in (c) for the matrix equation $\mathbf{AX} + \mathbf{BX} = \mathbf{B}$.

28. (Optional) If $\mathbf{A}^{-1}$ exists and $\mathbf{A}^2 + 3\,\mathbf{A} - 2\,\mathbf{I} = \mathbf{0}$, find $\mathbf{A}^{-1}$.

29. Given the system of linear equations
$$\begin{aligned} x + 2y &= 5 \\ 3x - 4y &= 6 \end{aligned}$$

(a) Write the system as a matrix equation of the form $\mathbf{AX} = \mathbf{B}$.

(b) Solve for $\mathbf{X}$ using $\mathbf{A}^{-1}$.

30. Determine the rank of each of the following matrices by reducing them to echelon form:

(a) $\begin{bmatrix} 1 & 5 & -1 \\ -2 & -10 & 2 \\ 4 & 8 & -4 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 4 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 4 & 6 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & -2 \\ -3 & 6 \\ 7 & -14 \end{bmatrix}$

31. Find the solution to each of the following systems of linear equations.

(a)
$$\begin{aligned} x + 4y - 2z &= 4 \\ 2x + 7y - z &= -2 \\ 2x + 9y - 7z &= 1 \end{aligned}$$

(b)
$$\begin{aligned} x + y \phantom{{}+8z} &= 0 \\ 4x + 6y + 8z &= 0 \\ 2y + 8z &= 0 \end{aligned}$$

(c)
$$\begin{aligned} x + y + z &= 0 \\ 2x + 5y + 3z &= 1 \\ -x + 2y + z &= 2 \end{aligned}$$

(d)
$$\begin{aligned} x_1 + x_3 &= 3 \\ x_2 + x_3 &= 6 \\ x_2 + x_4 &= 0 \\ x_1 + x_2 + x_3 + x_4 &= 2 \end{aligned}$$

(e)
$$\begin{aligned} x + 2y - z &= 5 \\ 2x - y + 3z &= 0 \end{aligned}$$

32. A system of linear equations with 5 unknowns has a unique solution.

(a) What is true about the ranks of the coefficient matrix and the augmented matrix?

(b) What must be true about the number of equations?

33. A system of linear equations with 6 unknowns has no solution. The rank of the coefficient matrix is 5.

(a) What is true about the rank of the augmented matrix?

(b) What must be true about the number of equations?

34. A system of 6 linear equations with 7 unknowns has many solutions. The rank of the coefficient matrix is 4.

(a) What is true about the rank of the augmented matrix?

(b) How many of the unknowns are "free" and can be assigned arbitrary parameters?

35. A system of homogeneous linear equations with 7 unknowns has only the trivial solution. What is the rank of the coefficient matrix?

36. Find the equation of the parabola $y = ax^2 + bx + c$ which passes through the points (1,1), (2,7), and (-1, -5). (Hint: when is a point on a curve?)

37. The <u>trace</u> of an nxn matrix $\mathbf{A} = [a_{ij}]$ is defined to be the sum of the diagonal elements, i.e. $\operatorname{tr}\mathbf{A} = \sum_{i=1}^{n} a_{ii}$. For example, if $\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$, then $\operatorname{tr}\mathbf{A} = 2+5 = 7$.

For the matrices $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} -3 & 2 \\ -1 & 1 \end{bmatrix}$

verify the following:

(a) $\operatorname{tr} 3\mathbf{A} = 3 \operatorname{tr}\mathbf{A}$

(b) $\operatorname{tr}(\mathbf{A}+\mathbf{B}) = \operatorname{tr}\mathbf{A} + \operatorname{tr}\mathbf{B}$

(c) $\operatorname{tr}\mathbf{AB} = \operatorname{tr}\mathbf{BA}$

(d) $\operatorname{tr}(\mathbf{B}^{-1}\mathbf{AB}) = \operatorname{tr}\mathbf{A}$

38. (a) Show that the vector $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$ is a linear combination of the vectors

$\begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

(b) Show that the vector $\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ is not a linear combination of the vectors

$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

39. (a) Show that the vectors $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$, and $\begin{bmatrix} 4 \\ 4 \\ 5 \end{bmatrix}$ are linearly dependent.

(b) Show that the vectors $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent.

40. Determine whether the vectors $\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$

are linearly dependent or linearly independent.

41. [Optional; do if you want to see the matrix form of the normal equations]

In a study of the possible relationship between smoking and lung cancer, the variable x represents cigarette consumption (per capita) in 1930 and y represents lung cancer deaths (per million) in 1950. Note: the study included 11 countries but only 4 are given here to simplify the computation.

	x	y
Norway	250	96
Sweden	265	115
Denmark	375	160
Holland	480	245

Use the matrix form of the normal equations to obtain the parameters in the linear relationship $y = b_0 + b_1x$.

42. Consider the matrices $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}$. Find

(a) det $\mathbf{A}$ (b) det $\mathbf{B}$ (c) det $\mathbf{AB}$ (d) det $\mathbf{BA}$ (e) det $\mathbf{A}^{-1}$ (f) det $\mathbf{A}'$

43. Given the matrix $\mathbf{A} = \begin{bmatrix} -1 & 4 & 5 \\ 3 & 6 & 1 \\ 2 & 5 & 7 \end{bmatrix}$

(a) Find det $\mathbf{A}$ by expanding by cofactors along the first row.

(b) Find det $\mathbf{A}$ by expanding by cofactors along the second column.

44. Find the determinant of each of the following matrices:

(a) $\begin{bmatrix} 1 & -1 & -3 \\ 4 & 2 & 2 \\ 2 & 1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ (c) $\begin{bmatrix} 1-x & 2 & -1 \\ 0 & -x & 1 \\ 3 & 0 & 2-x \end{bmatrix}$

45. If $\mathbf{A}$ and $\mathbf{B}$ are arbitrary nxn matrices, which of the following statements are always true? If a statement is not always true, give an example to illustrate when it could be false.

(a) det $(\mathbf{A} + \mathbf{B})$ = det $\mathbf{A}$ + det $\mathbf{B}$

(b) det $\mathbf{AB}$ = (det $\mathbf{A}$)(det $\mathbf{B}$)

(c) det 3 $\mathbf{A}$ = 3 det $\mathbf{A}$

46. Use the adjoint method to find the inverse of each of the following matrices:

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

47. Use Cramer's Rule to solve the system

$$\begin{aligned} x + 2y &= 5 \\ 3x - 4y &= 6 \end{aligned}$$

48. Find the eigenvalues and the eigenvectors of the following matrices:

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

(d) For each of the matrices **M** in (a), (b), (c), show that the sum of the eigenvalues equals the trace, i.e. tr $\mathbf{M} = \sum \lambda_i$ and that the product of the eigenvalues equals the determinant, i.e. det $\mathbf{M} = \prod \lambda_i$.

(e) [Optional: see Namboodiri pp.83-86] For each of the matrices **M** above, obtain the diagonalization $\mathbf{M} = \mathbf{P}\ \Lambda\ \mathbf{P}^{-1}$ (that is, determine **P** and Λ).

49. [Optional; for those who want to see principal components]

Suppose that the covariance matrix for two variables X_1 and X_2 is $\begin{bmatrix} 14 & 6 \\ 6 & 5 \end{bmatrix}$

(a) Show that the eigenvalues of the matrix are 17 and 2.

(b) Show that the first principal component is $2/\sqrt{5}\,X_1 + 1/\sqrt{5}\,X_2$, which gives almost 90% of the total variance.

50. [Optional] Suppose that the covariance matrix of a study of three variables X_1, X_2, and X_3 is

$$\mathbf{C} = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

(a) Show that the eigenvalues of **C**, arranged from largest to smallest are 6, 3, and 1.

(b) Find the first principal component and determine what percent of the total variance is accounted for by this new variable.

(c) Find the second principal component and determine what percent of the total variance is accounted for by this new variable. The first two principal components account for how much of the total variance?

ANSWERS - MATRIX ALGEBRA

1. (a) 6,000 (b) 43,000 (c) 1,189,000 (d) Metro Toronto and York
   (e) 385,000 (f) 740,000

2. (a) Less than 1% $\left(\frac{8}{4304} \times 100 \simeq 0.19\%\right)$ (b) About 40% $\left(\frac{448+259+218+282+533}{4304} \times 100\right)$
   (c) The richest (d) About 40% $\left(\frac{1711}{4304} \times 100\right)$

3. $a = 12,\ b = -6,\ c = 2$

4. (a) $\begin{bmatrix} 3 & 1 \\ 3 & 7 \\ 1 & 11 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ (d) $\begin{bmatrix} 3 & 1 \\ 3 & 7 \\ 1 & 11 \end{bmatrix}$ (e) B (f) $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

5. $\begin{bmatrix} 96.7 & -7.7 & 8.7 \\ -16.1 & -135.8 & -55.8 \\ -0.3 & 117.0 & 0.1 \\ 48.6 & -457.9 & -65.9 \end{bmatrix}$ The entries give the increase (+) or decrease (−) in energy production over energy consumption in each type (columns) by the four countries (rows).

6. (a) $\begin{bmatrix} 3 & 9 & -6 \\ -3 & 6 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 7 & -7 & 12 \\ -5 & 10 & -12 \end{bmatrix}$ (c) $\begin{bmatrix} -2.5 & 6.5 & -8.0 \\ 1.5 & -3.0 & 6.0 \end{bmatrix}$

   (d) $X = B - A = \begin{bmatrix} 2 & -8 & 9 \\ -1 & 2 & -6 \end{bmatrix}$ (e) $2X = B - A$ so $X = \frac{1}{2}(B-A)$
   $\therefore X = \begin{bmatrix} 1 & -4 & 4.5 \\ -0.5 & 1 & -3 \end{bmatrix}$

7. Calculate both sides of each of the equations. For example in (b):

$$5(A+B) = 5\begin{bmatrix} 4 & -2 & 5 \\ -3 & 6 & -6 \end{bmatrix} = \begin{bmatrix} 20 & -10 & 25 \\ -15 & 30 & -30 \end{bmatrix}$$

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

$$= \begin{bmatrix} 5 & 15 & -10 \\ -5 & 10 & 0 \end{bmatrix} + \begin{bmatrix} 15 & -25 & 35 \\ -10 & 20 & -30 \end{bmatrix}$$

$$= \begin{bmatrix} 20 & -10 & 25 \\ -15 & 30 & -30 \end{bmatrix}$$

$$\therefore \quad 5(A+B) = 5A + 5B$$

8. (a) $\begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}$ (b) $\begin{bmatrix} 8 \\ 2 \\ 3 \end{bmatrix}$ (c) $\begin{bmatrix} -1 \\ -4 \\ 4 \end{bmatrix}$ (d) -5 (e) 14 (f) $x = 5$

(g) $|b| = \sqrt{b' \cdot b} = \sqrt{2^2+2^2+(-1)^2} = \sqrt{9} = 3$

(h) $u = \frac{1}{|b|} b = \frac{1}{3} b = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix}$ check: $|u| = 1$

9. (a) $[8]$ (b) $\begin{bmatrix} 8 \\ 5 \end{bmatrix}$ (c) $\begin{bmatrix} 3 \\ 8 \end{bmatrix}$ (d) $[-11]$ (e) $[5 \ \ 6]$ (f) $\begin{bmatrix} 5 \\ 11 \end{bmatrix}$

10. (a) $\begin{bmatrix} 5 & 14 \\ 18 & 39 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 14 & 16 \\ 2 & 19 & 20 \\ 3 & 24 & 24 \end{bmatrix}$ (c) $\begin{bmatrix} -1 & 4 & 8 \\ 3 & -6 & -16 \end{bmatrix}$ (d) not defined

11. (a) $\begin{bmatrix} 6 & -3 & -3 \\ -3 & -3 & 6 \\ -3 & 6 & -3 \end{bmatrix}$ (b) Same as (a) (c) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (d) Same as (c)

(e) $\begin{bmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{bmatrix}$ (f) B (g) 0

12. $AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

$= \frac{1}{ad-bc} \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

Similarly, $BA = I$. Later we will say $B = A^{-1}$. (the inverse of A)

13. Calculate both sides of each of the equations.

14. (a), (c), and (d) are false. For example, let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ (other examples are possible)

(b) is true. $(A-B)^2 = (A-B)(A-B) = A(A-B) - B(A-B)$
$= A^2 - AB - BA + B^2 = A^2 + B^2 - AB - BA$

15. (a) $a = -1$, $c = 5$, b any number. (b) $a = -2$, $b = 2$, $c = 0$

16. (a) There are 11,600 registered Democrats and 8,400 registered Republicans.

(b) The Democrats would win with about 58% of the votes.

(c) There would be 10,100 registered Republicans and 9,900 registered Democrats. The Republicans would be expected to win.

17. (a) $PM \approx [12.54 \quad 87.46]$

(b) $PM^2 = (PM)M \approx [12.47 \quad 87.53]$

18. (a) $\begin{bmatrix} 2 & -3 & 4 \\ 3 & 4 & 2 \\ 4 & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 0 & -3 \\ 3 & -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 1 \\ 3 & -5 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$

19. $y = \begin{bmatrix} 60.4 \\ 79.4 \\ 77.6 \\ 86.0 \\ 95.3 \\ 95.9 \end{bmatrix}$, $X = \begin{bmatrix} 1 & 0 \\ 1 & 11 \\ 1 & 20 \\ 1 & 30 \\ 1 & 41 \\ 1 & 47 \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, $e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix}$

20. (a) $x^2-4xy+3y^2+2yz-4z^2$

(b) $4x^2+6xy+y^2-8xz-z^2 = 4x^2+3xy+3yx+y^2-4xz-4zx-z^2$

Let $X=[x \quad y \quad z]$ and $A=\begin{bmatrix} 4 & 3 & -4 \\ 3 & 1 & 0 \\ -4 & 0 & -1 \end{bmatrix}$

21.
$$A=\begin{array}{c} \begin{array}{cc} X_1 & X_2 \end{array} \\ \begin{bmatrix} 24 & 52 \\ 23 & 43 \\ 17 & 43 \\ 13 & 19 \\ 26 & 35 \\ 25 & 64 \\ 10 & 19 \\ 13 & 24 \\ 10 & 42 \\ 13 & 48 \end{bmatrix} \end{array}$$

$$A'A = \begin{bmatrix} 3402 & 7271 \\ 7271 & 17089 \end{bmatrix}$$

$$A'UA = \begin{bmatrix} 30276 & 67686 \\ 67686 & 151321 \end{bmatrix}$$

$$S = A'A - \tfrac{1}{10}A'UA = \begin{bmatrix} 374.4 & 502.4 \\ 502.4 & 1956.9 \end{bmatrix}$$

$$\therefore C = \tfrac{1}{9}S = \begin{bmatrix} 41.6 & 55.82 \\ 55.82 & 217.43 \end{bmatrix}$$

22. If $B=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $AB=I$ then we have $a+3c=1$ and $2a+6c=0$

i.e. we have the system of equations $\begin{cases} a+3c=1 \\ a+3c=0 \end{cases}$

which has no solution.

23.

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

(c) $\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[-2R_1+R_3]{R_1+R_2} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & -2 & 0 & 1 \end{array}\right]$

$\xrightarrow{-R_2+R_3} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & -3 & -1 & 1 \end{array}\right] \xrightarrow{-R_3} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 3 & 1 & -1 \end{array}\right]$

$\xrightarrow[-R_3+R_1]{R_2+R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 3 & 1 & -1 \end{array}\right]$ $\therefore$ The inverse is $\begin{bmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \\ 3 & 1 & -1 \end{bmatrix}$

(d) The inverse does not exist.

24. (a) $B = A'$ so $B^{-1} = (A')^{-1} = (A^{-1})' = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}$

(b) $\frac{1}{14}\begin{bmatrix} 1 & 3 \\ 4 & -2 \end{bmatrix}$ (c) $(AC)^{-1} = C^{-1}A^{-1}$ (d) Calculate both sides.

25. $AA' = \begin{bmatrix} x & -2x \\ 2x & x \end{bmatrix}\begin{bmatrix} x & 2x \\ -2x & x \end{bmatrix} = \begin{bmatrix} 5x^2 & 0 \\ 0 & 5x^2 \end{bmatrix}$

$\therefore$ If $AA' = I$ then $5x^2 = 1$ so $x = \pm \frac{1}{\sqrt{5}}$

26. $AA' = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}\begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} \cos^2\phi+\sin^2\phi & 0 \\ 0 & \sin^2\phi+\cos^2\phi \end{bmatrix}$

But $\sin^2\phi + \cos^2\phi = 1$ so $AA' = I$ $\therefore$ $A^{-1} = A' = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$

27. (a) $B = A^{-1}C$ (b) $A = P^{-1} B Q^{-1}$

(c) $AX = B + 2X \Rightarrow AX - 2X = B \Rightarrow AX - 2IX = B$

$\Rightarrow (A-2I)X = B$.

We need A to be a square matrix, say A is $m \times m$. If B is an $m \times m$ matrix and $A - 2I$ is invertible, then $X = (A-2I)^{-1}B$.

(d) If A and B are $m \times m$ and $A+B$ is invertible, then $X = (A+B)^{-1}B$.

28. $A^2 + 3A - 2I = 0 \Rightarrow 2I = A^2 + 3A \Rightarrow I = \frac{1}{2}(A^2 + 3A)$

$\Rightarrow I = \frac{1}{2}(A^2 + 3IA) \Rightarrow I = \frac{1}{2}(A+3I)A$

$\Rightarrow IA^{-1} = \frac{1}{2}(A+3I)AA^{-1} \Rightarrow A^{-1} = \frac{1}{2}(A+3I)$

29. (a) $\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}^{-1} = \frac{1}{-10}\begin{bmatrix} -4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.2 \\ 0.3 & -0.1 \end{bmatrix}$

$\therefore \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.4 & 0.2 \\ 0.3 & -0.1 \end{bmatrix}\begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3.2 \\ 0.9 \end{bmatrix}$

30. (a) 2 (b) 4 (c) 1

31. (a) $\left[\begin{array}{ccc|c} 1 & 4 & -2 & 4 \\ 2 & 7 & -1 & -2 \\ 2 & 9 & -7 & 1 \end{array}\right] \xrightarrow[-2R_1+R_3]{-2R_1+R_2} \left[\begin{array}{ccc|c} 1 & 4 & -2 & 4 \\ 0 & -1 & 3 & -10 \\ 0 & 1 & -3 & -7 \end{array}\right]$.

$\xrightarrow{R_2+R_3} \left[\begin{array}{ccc|c} 1 & 4 & -2 & 4 \\ 0 & -1 & 3 & -10 \\ 0 & 0 & 0 & -17 \end{array}\right]$

The last row (equation) reads: $0x + 0y + 0z = -17$, which is impossible for any choice of x, y, z. So the system has no solution.

(b) $\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 4 & 6 & 8 & 0 \\ 0 & 2 & 8 & 0 \end{array}\right] \xrightarrow{-4R_1+R_2} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 2 & 8 & 0 \\ 0 & 2 & 8 & 0 \end{array}\right] \xrightarrow{-R_2+R_3} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 2 & 8 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

$\xrightarrow{\frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} \underline{1} & 0 & -4 & 0 \\ 0 & \underline{1} & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

Let $z = a$ (arbitrary)
$y = -4a$
$x = 4a$

(c) $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 5 & 3 & 1 \\ -1 & 2 & 1 & 2 \end{array}\right] \xrightarrow[R_1+R_3]{-2R_1+R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 3 & 1 & 1 \\ 0 & 3 & 2 & 2 \end{array}\right] \xrightarrow{\frac{1}{3}R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 3 & 2 & 2 \end{array}\right]$

$\xrightarrow{-3R_2+R_3} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow[-R_3+R_1]{-\frac{1}{3}R_3+R_2} \left[\begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]$

$\xrightarrow{-R_2+R_1} \left[\begin{array}{ccc|c} \underline{1} & 0 & 0 & -1 \\ 0 & \underline{1} & 0 & 0 \\ 0 & 0 & \underline{1} & 1 \end{array}\right]$ $\therefore x=-1,\ y=0,\ z=1$

(d) No solution (e) $x = 1-a,\ y = 2+a,\ z = a$ (any number)

32. (a) They are both 5 (b) There are at least 5

33. (a) It is 6 (b) There are at least 6

34. (a) It is 4 (b) There are 3

35. It is 7.

36. The point $(1,1)$ is on the curve so $a(1)^2 + b(1) + c = 1$. Similarly for the other points. We obtain the system of equations

$$\begin{cases} a + b + c = 1 \\ 4a + 2b + c = 7 \\ a - b + c = -5 \end{cases}$$

The solution is $a = 1,\ b = 3,\ c = -3$ so $y = x^2 + 3x - 3$

37. Calculate both sides of each equation. For example, in (b) we have

$$A+B=\begin{bmatrix}-2 & 4\\ 2 & 5\end{bmatrix} \quad \text{so} \quad tr(A+B) = -2+5 = 3$$

$$tr A = 1+4 \quad \text{and} \quad tr B = -3+1 \quad \text{so} \quad tr A + tr B = 5-2 = 3$$

38. (a) $a\begin{bmatrix}3\\5\end{bmatrix}+b\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}-2\\3\end{bmatrix} \Rightarrow \begin{cases}3a+2b=-2\\ 5a+b=3\end{cases}$

$$\therefore \left[\begin{array}{cc:c}3 & 2 & -2\\ 5 & 1 & 3\end{array}\right] \to \cdots \to \left[\begin{array}{cc:c}1 & 0 & \frac{8}{7}\\ 0 & 1 & -\frac{19}{7}\end{array}\right]$$

$$\therefore \begin{bmatrix}-2\\3\end{bmatrix} = \frac{8}{7}\begin{bmatrix}3\\5\end{bmatrix} - \frac{19}{7}\begin{bmatrix}2\\1\end{bmatrix}$$

(b) The system $\left[\begin{array}{ccc:c}1 & -1 & 0 & 1\\ 2 & 0 & 1 & -2\\ 3 & 1 & 2 & 3\end{array}\right]$ has no solution.

39. (a) $A=\begin{bmatrix}2 & -2 & 4\\ 1 & 1 & 4\\ 1 & 2 & 5\end{bmatrix}$ leads to the echelon form $\begin{bmatrix}1 & 1 & 4\\ 0 & 1 & 1\\ 0 & 0 & 0\end{bmatrix}$

(b) $A=\begin{bmatrix}2 & -2 & 1\\ 1 & 1 & 0\\ 1 & 2 & 1\end{bmatrix}$ leads to the echelon form $\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$

40. Linearly dependent.

41. $X=\begin{bmatrix}1 & 250\\ 1 & 265\\ 1 & 375\\ 1 & 480\end{bmatrix}, \quad Y=\begin{bmatrix}96\\115\\160\\245\end{bmatrix} \quad \therefore X'X=\begin{bmatrix}4 & 1370\\ 1370 & 503750\end{bmatrix}$

$$(X'X)^{-1} = \frac{1}{138100}\begin{bmatrix}503750 & -1370\\ -1370 & 4\end{bmatrix}$$ e.g. see problem #12

$$X'Y = \begin{bmatrix}616\\232075\end{bmatrix}$$

$$\therefore \hat{\beta} = (X'X)^{-1}(X'Y) = \begin{bmatrix}-55.27\\0.61\end{bmatrix} \quad \text{so} \quad \hat{y} = 0.61x - 55.27$$

42. (a) -2 (b) 11 (c) -22 (d) -22 (e) $-\frac{1}{2}$ (f) -2

43. (a) $-1(1)(42-5)+4(-1)(21-2)+5(1)(15-12) = -98$
(b) $4(-1)(21-2)+6(1)(-7-10)+5(-1)(-1-15) = -98$

44. (a)
$$A = \begin{bmatrix} 1 & -1 & -3 \\ 4 & 2 & 2 \\ 2 & 1 & -1 \end{bmatrix} \xrightarrow[-2R_1+R_3]{-4R_1+R_2} \begin{bmatrix} 1 & -1 & -3 \\ 0 & 6 & 14 \\ 0 & 3 & 5 \end{bmatrix} \xrightarrow[\ominus]{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & -1 & -3 \\ 0 & 3 & 5 \\ 0 & 6 & 14 \end{bmatrix}$$
$$\xrightarrow{-2R_2+R_3} \begin{bmatrix} 1 & -1 & -3 \\ 0 & 3 & 5 \\ 0 & 0 & 4 \end{bmatrix} = B \qquad \therefore \det A = -\det B = -(1)(3)(4) = -12.$$

(b) 0 (c) Expand by a row or a column to get $-x^3+3x^2-5x+6$

45. (a) and (c) are false. For example, use $A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

46. (a) $\det A = -2$; $\operatorname{adj} A = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$ $\therefore A^{-1} = \frac{1}{-2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$

(see Problem 23 (a))

(b) $A^{-1} = \frac{1}{-5}\begin{bmatrix} -2 & 1 & -1 \\ 1 & 2 & -2 \\ 2 & -1 & -4 \end{bmatrix}$

47. $x = \frac{-32}{-10} = 3.2$; $y = \frac{-9}{-10} = 0.9$

48. (a) $\lambda = 1$, $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$; $\lambda = 7$, $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

(b) $\lambda = 0$, $\begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}$; $\lambda = 2$, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$; $\lambda = -3$, $\begin{bmatrix} -4 \\ 5 \\ 0 \end{bmatrix}$

(c) $A-\lambda I = \begin{bmatrix} 2-\lambda & 2 & 0 \\ 1 & -1-\lambda & 2 \\ 0 & 1 & 2-\lambda \end{bmatrix}$

$\det(A-\lambda I) = (2-\lambda)\begin{vmatrix} -1-\lambda & 2 \\ 1 & 2-\lambda \end{vmatrix} - 1\begin{vmatrix} 2 & 0 \\ 1 & 2-\lambda \end{vmatrix}$ from expanding down the first column

$= (2-\lambda)(\lambda^2-\lambda-4) - 2(2-\lambda) = (2-\lambda)(\lambda^2-\lambda-6)$
$= (2-\lambda)(\lambda+2)(\lambda-3)$

$\therefore$ The eigenvalues are $\lambda=2$, $\lambda=-2$, and $\lambda=3$

For $\lambda=2$, solving $\left[\begin{array}{ccc|c} 0 & 2 & 0 & 0 \\ 1 & -3 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

we get the eigenvector $\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ (and any scalar multiple of it).

Similarly, for $\lambda=-2$ we get the eigenvector $\begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix}$ and for $\lambda=3$ we get $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$

(d) (a) $\operatorname{tr} M = 8$ and $\lambda_1+\lambda_2 = 1+7 = 8$; $\det M = 7$ and $\lambda_1\lambda_2 = 1\cdot 7 = 7$
(b) $\operatorname{tr} M = -1$ and $\lambda_1+\lambda_2+\lambda_3 = -1$; $\det M = 0$ and $\lambda_1\lambda_2\lambda_3 = 0$
(c) $\operatorname{tr} M = 3$ and $\lambda_1+\lambda_2+\lambda_3 = 3$; $\det M = -12$ and $\lambda_1\lambda_2\lambda_3 = (2)(-2)(3) = -12$

(e) (a) $P = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$, where the columns are the eigenvectors

$\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$, where the diagonal elements are the eigenvalues.

Note: the order of the eigenvalues on the diagonal of Λ has to correspond to the order of the eigenvectors in P

(b) $P = \begin{bmatrix} -1 & 1 & -4 \\ -1 & 0 & 5 \\ 3 & 0 & 0 \end{bmatrix}$; $\Lambda = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$

(c) $P = \begin{bmatrix} 2 & -2 & 2 \\ -4 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$; $\Lambda = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

49. (a) $\det(A-\lambda I) = \begin{vmatrix} 14-\lambda & 6 \\ 6 & 5-\lambda \end{vmatrix} = \lambda^2 - 19\lambda + 34 = (\lambda-17)(\lambda-2)$ $\therefore \lambda = 17$ and $\lambda = 2$.

(b) The eigenvector corresponding to $\lambda=17$ is $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ (or any scalar multiple of it)

a unit vector in the direction of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is $\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$

$\therefore$ The first principal component is $\frac{2}{\sqrt{5}}X_1 + \frac{1}{\sqrt{5}}X_2$

The total variance is $14+5=19$. The first principal component accounts for $\frac{17}{19} \approx 0.895$, or almost 90% of the total variance.

50.

(a) $C - \lambda I = \begin{bmatrix} 4-\lambda & 1 & 1 \\ 1 & 3-\lambda & 2 \\ 1 & 2 & 3-\lambda \end{bmatrix}$. Expand down first column to get $\det(C - \lambda I)$. You could do other rows or columns.

$$(4-\lambda)\begin{vmatrix} 3-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ 2 & 3-\lambda \end{vmatrix} + 1\begin{vmatrix} 1 & 1 \\ 3-\lambda & 2 \end{vmatrix}$$

$$= (4-\lambda)(\lambda^2 - 6\lambda + 5) - (1-\lambda) + (\lambda - 1)$$
$$= (4-\lambda)(\lambda-5)(\lambda-1) + (\lambda-1) + (\lambda-1)$$
$$= (\lambda-1)\left[(4-\lambda)(\lambda-5) + 1 + 1\right] = (\lambda-1)(-\lambda^2+9\lambda-18)$$
$$= -(\lambda-1)(\lambda^2 - 9\lambda + 18)$$
$$= -(\lambda-1)(\lambda-3)(\lambda-6)$$

$\therefore$ The eigenvalues are 6, 3, 1.

The eigenvector corresponding to $\lambda = 6$ is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, or a scalar multiple of it

a unit vector in the same direction is $\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ (see Problem 8(h))

(b) The first principal component is $Y_1 = \frac{1}{\sqrt{3}}X_1 + \frac{1}{\sqrt{3}}X_2 + \frac{1}{\sqrt{3}}X_3$

The total variance $= \operatorname{Tr} C = 10$

The first principal component accounts for $\frac{6}{10}$ of this total or 60%

(c) The eigenvector corresponding to $\lambda = 3$ is $\begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$ or, a unit vector in the same direction is $\frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$

The second principal component is $Y_2 = \frac{2}{\sqrt{6}}X_1 - \frac{1}{\sqrt{6}}X_2 - \frac{1}{\sqrt{6}}X_3$

The second principal component accounts for $\frac{3}{10}$ or 30% of the total variance.

Y_1 and Y_2 together account for $\frac{6+3}{10}$ or 90% of the total variance

EXAMPLES OF NON-LINEAR FUNCTIONS

1. U.S. Military Spending (Center for Defense Information; outlays in billions of 2002 dollars)

Year	1980	1982	1984	1986	1988	1990	1992	1994	1996	1998
Amount ($billion)	303.4	339.4	381.7	426.6	426.4	409.7	379.5	338.6	307.4	296.7

In plotting the data, a parabola is suggested. The best fitting *quadratic* is

$$y = 306.7 + 25.1x - 1.5x^2.$$ Note: $x = 0$ corresponds to 1980.

2. AIDS Cases in the United States (Centers for Disease Control data subject to substantial retrospective changes. AIDS was not a notifiable disease until 1984)

Year	1984	1986	1988	1990	1992	1994	1996	1998	2000	2002
Cumulative Total (x1000)	11	43	108	201	341	494	625	718	800	883

A study of the data (as well as theoretical considerations) indicates that a cubic equation is an appropriate model. The best fitting *cubic* (with rounding) is

$$y = - 0.25\, x^3 + 7.14\, x^2 + 0.34x + 10.69.$$ Note: $x = 0$ corresponds to 1984.

3. "Cube Law" (Readings in Math.Soc.Sci., Lazarsfeld & Henry; also Amer. Pol. Sci. Review, June 1986)

In a two-party system, let y be the proportion of seats won in an election by one party and let x be the proportion of votes won in the election by that party. Then we have the *rational* function

$$y = \frac{x^3}{(1-x)^3 + x^3}$$

4. Prisoners (data from the U.S. Department of Justice)

For the years 1970-2003 the number of sentenced prisoners (per 100,000 residents) under jurisdiction of State and Federal correctional authorities can be approximated by the *exponential* function $P(t) = 85.5\,(1.06)^t$, where $t = 0$ corresponds to 1970.

5. Urban Concentration (data from U.S. Census Bureau)

The most populous 100 cities in 2000 are ranked and their populations are noted. The population P (in millions) in terms of the rank R is given approximately by the *power* function

$$P = 5.79\, R^{-0.74}$$ (e.g. San Antonio has R = 9 and P = 1.15; the formula gives P = 1.14)

6. Daylight in Ann Arbor (data from U.S. Naval Observatory)

The number of minutes of daylight in Ann Arbor is approximated by the *trigonometric* function

$$f(x) = 182 \sin(0.017x - 1.39) + 728,$$ where x denotes the day of the year.

USING DERIVATIVES FOR THE GRAPH OF A FUNCTION

Given a function $y = f(x)$ and suppose that we have found $f'(x)$ and $f''(x)$.

1. Some information from the first derivative f'

   If $f'(a) > 0$ then f is increasing at $x = a$.
   If $f'(a) < 0$ then f is decreasing at $x = a$.
   If $f'(a) = 0$ then f is stationary at $x = a$.

2. Some information from the second derivative f''

   If $f''(a) > 0$ then f is concave up at $x = a$.
   If $f''(a) < 0$ then f is concave down at $x = a$.
   If $f''(a) = 0$ then we will need to examine further the situation at $x = a$ (see the examples below).

3. Local Maxima and Minima (Extrema)

   Suppose that f is stationary at $x = a$ (i.e. $f'(a) = 0$)

   (a) First derivative test:

   - If f' is negative to the left of a and positive to the right of a then f has a local minimum at $x = a$.
   - If f' is positive to the left of a and negative to the right of a then f has a local maximum at $x = a$.

   (b) Second derivative test:

   - If $f''(a) > 0$ then f has a local minimum at $x = a$.
   - If $f''(a) < 0$ then f has a local maximum at $x = a$.
   - If $f''(a) = 0$ then this test fails and we must use the first derivative test (see the examples below).

4. Inflection Point

   If f'' is positive on one side of $x = a$ and negative on the other side of $x = a$ then f has an *inflection point* at $x = a$.

   This often (but not always) happens when $f''(a) = 0$; you must test on either side of $x = a$ to see if there is a change in the concavity.

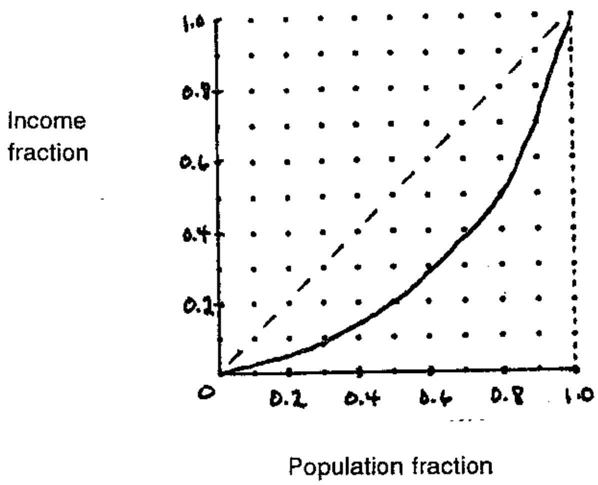
Examples: Compare the situation for the functions $f(x) = x^3$, $f(x) = x^4$, and $f(x) = - x^6$. All three have $f'(0) = 0$ and $f''(0) = 0$ but the first has an inflection point at $x = 0$, the second has a local minimum at $x = 0$, and the third has a local maximum at $x = 0$.

Lorenz Curves -- The Gini Index

If we consider the population of the United States and the income each person receives, we might note some inequality. For example, the U.S. Census Bureau reports that in 2000 the lowest 20% of the population shared about 4% of the total household income while the top 20% of the population shared about 50% of the total income. If we graph the fraction of income against the fraction of the population and draw a smooth curve through the data points, we obtain what is called a Lorenz curve. This type of curve is also obtained if we graph the fraction of the population of New York State in 1960 against the fraction of seats represented by that fraction in the State Assembly. And we obtain this type of curve if we graph the fraction of the countries of the world against the fraction of the total energy used. There are other examples as well.

Let's return to the original example (household income and population of the U.S. in 2000).

Fraction of Population	Fraction of Income
0.00	0.00
0.20	0.04
0.40	0.13
0.60	0.27
0.80	0.50
1.00	1.00



Note: If there is absolute equality then the Lorenz curve would look like the dashed curve (- - -) and if there is absolute inequality it would look like the dotted curve (····). Usually it looks like something in between like the solid curve (—).

One of the measures of inequality is the Gini Index (GI). It is obtained by comparing the area between the Lorenz curve and the dashed line with the area of the right triangle formed by the dashed line, the dotted line, and the horizontal axis.

$$GI = \frac{\text{area between the Lorenz curve and the dashed line}}{\text{area of the right triangle}} = \frac{\text{area between curve and dashed line}}{1/2}$$

$$= 2\,(\text{area between the Lorenz curve and the dashed line})$$
$$= 2\,(1/2 \;-\; \text{area under the Lorenz curve})$$

So **GI = 1 - 2 A** , where A is the area under the Lorenz curve

Note: For absolute equality, GI = 1 - 2(1/2) = 0 and for absolute inequality, GI = 1 - 2(0) = 1.

PROBLEM: How do we find A, the area under a curve?

MATHEMATICS FOR SOCIAL SCIENTISTS II

CALCULUS EXERCISES

1. The U.S. Census Bureau reported that the median age for women at first marriage was 22.0 in 1980 and 25.1 in 2000. Suppose that there is a linear relationship between the median age y and time t (let $t = 0$ correspond to 1980).

   (a) Obtain the linear relationship.

   (b) Write a sentence that explains the meaning of the slope and the y-intercept in the linear relationship you obtained in (a).

2. The National Highway Traffic Safety Administration reports fatalities that are alcohol-related. From 1982-2000 the percentage is approximately a linear function of time. In 1982 the percentage was 57% and declined approximately 0.8% a year. If the linear pattern would continue, what would be the best estimate for the percentage in the year 2002?

3. The table below shows the number per thousand resident population of sentenced prisoners under jurisdiction of State and Federal correctional authorities for selected years. Source: U.S. Department of Justice

Year	1960	1964	1970	1975	1985	1995	2002
Number	117	111	96	111	202	411	476

   Find the average rate of change for the number per thousand between the years

   (a) 1960 and 2002 (b) 1960 and 1970 (c) 1964 and 1975

   (d) 1970 and 2002 (e) 1960 and 1975

4. Let $y = f(x) = x^2 - 4x + 7$. Find the average rate of change of y as x changes from

   (a) 1 to 4 (b) 1 to 3 (c) 1 to 2

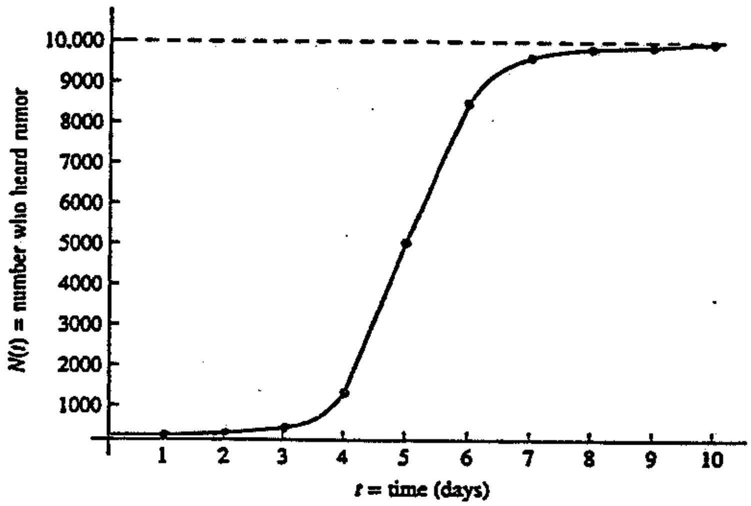
5. For $y = f(x) = x^2 - 4x + 7$ and $x = 1$ a fixed value of x,

(a) Find $\frac{\Delta y}{\Delta x}$ when $\Delta x = 3$; when $\Delta x = 2$; when $\Delta x = 1$; when $\Delta x = 0.1$; when $\Delta x = 0.01$.

(b) What do you think happens as we let Δx get smaller and smaller?

6. There is a mathematical model used by sociologists to study how a rumor spreads (see J. Coleman, *An Introduction to Mathematical Sociology).* Suppose there is a town of 10,000 people and N(t) denotes the number of people in the town who have heard a rumor after t days. Values of N(t) are tabulated and a smooth curve is drawn through the data points. For example, N(2) = 40 so after 2 days 40 people have heard the rumor.

t	N(t)
0	1
1	6
2	40
3	245
4	1368
5	5000
6	8631
7	9754
8	9960
9	9994
10	9999



(a) Use the given table to estimate the value of the derivative of this function at $t = 2$, $t = 5$, and $t = 8$.

(b) When does the rumor seem to spread less rapidly? Can you give a reason why this is so?

(c) When does the rumor seem to be spreading most rapidly? Can you give a reason why this is so?

7. Find $\frac{dy}{dx}$ for each of the following functions:

(a) $y = -25$ (b) $y = 15 - (3/2)\,x$ (c) $y = -4x^2$

(d) $y = x^{5/3}$ (e) $y = \frac{8}{x^5}$

8. Given that $f(x) = 3x^4 - 6x^2 + 7$

(a) Is f increasing or decreasing at $x = 2$?

(b) What is the slope of the tangent to the curve at $x = -2$?

(c) When is f stationary?

9. If $f(x) = 2x^3 - 3x^2 - 12x + 5$, for what values of x is f decreasing? (Give an interval of values that x can take on so that f is decreasing on that interval.)

10. Determine $\frac{dy}{dx}$ for each of the following:

(a) $y = (x^2 - 1)(3x + 10)$ (b) $y = \frac{2x}{5-2x^2}$

(c) $y = (2x - 5)^3$ (d) $y = \sqrt{x^5 - 3}$

11. Find the inflection points for each of the following functions:

(a) $f(x) = 2x^3 - 9x^2 + 12x$ (b) $f(x) = x^4 - 6x^2 + 5x - 6$

(c) $f(x) = 4x^3 - 3x^4$ (d) $f(x) = x^4 - 4x - 1$

12. Find the maximum and/or the minimum value of f(x) for each of the following:

(a) $f(x) = 10 + 24x - 3x^2$ (b) $f(x) = (1/3)\, x^3 - (1/2)\, x^2 - 6x$

(c) $f(x) = (1/4)\, x^4 - 8x^2 + 10$

13. Suppose that the annual profit P (in dollars) for a cable TV company obtained from the monthly fee r (in dollars) that the customers are charged is given by

$$P(r) = -20{,}000r^2 + 1{,}820{,}000r - 1{,}750{,}000$$

(a) What monthly fee would maximize the profit obtained from the monthly fees?

(b) What would be this maximum annual profit?

14. Climbing health care costs in the United States has been a source of concern for some time. Data from the U.S. Health Care Financing Administration shows the average yearly per-capita national health expenditures (in dollars) for selected calendar years as follows:

Year	Per-capita expenditures
1960	143
1965	204
1970	348
1975	591
1980	1067
1985	1761
1990	2738
1995	3698
2000	4637

(a) Plot the data. Do you think a linear function or a quadratic function would be more likely to model the data?

(b) Let x represent the year (x = 0 is 1960) and y represent the per-capita expenditures in dollars.

The best fit linear function is $y = 114.7x - 606.4$ and the best fit quadratic function is $y = 3.3x^2 - 15.4x + 152.6$. Use each to estimate per-capita expenditures in 2002.

15. Psychologists studying short term memory recall, found that the percentage P of recall was a function of the time t seconds after the item was last seen. They said that the data best fit the negative exponential function

$$P(t) = 16 + 84e^{-0.011t}.$$

(a) What is the percentage of recall after 60 seconds?

(b) How many seconds does it take for P to be 25%?

16. Determine the maximum and/or minimum of the function

$$y = f(x) = 8xe^{-2x} - 500$$

17. The growth of population in Mexico City proper was phenomenal during most of the 20th century, spurred by migration from the provinces and a high birthrate. From 1950 through 1970 it grew from 3,050,000 to 6,874,000.

(a) Let $P(t)$ be the population as a function of time and use the exponential model $P(t) = P_0 e^{rt}$, where $t = 0$ corresponds to 1950. Determine $P(t)$ explicitly (i.e. get P_0 and r).

(b) In 1970 the government began a concerted effort to reduce birthrates and to decrease migration through the creation of jobs in other regions. In 1990 the population was 8,236,000. What is the exponential growth rate for the years 1970-1990?

18. Someone determines that the cost (in dollars) of processing each application in a state welfare agency is given by

$$c(x) = 0.005x^2 - 16 \ln x + 70,$$

where x is the number of analysts working in the agency. Find the value of x that will minimize the cost for each application and give that cost.

19. Consider the function $f(x,y) = 20x^3 - 30y^3 + 10x^2y$

(a) Find $f(10,10)$

(b) Find the actual change in the value of f if x is increased by 1 to 11 and y remains 10, that is find

$$\frac{f(10 + 1, 10) - f(10,10)}{1}$$

(c) Find the instantaneous rate of change of f when x changes and y remains 10, that is, find the value of $\frac{\partial f}{\partial x}$ at (10,10).

(d) Modify (b) and (c) to study the case where x remains 10 and y is increased by 1 and the value of $\frac{\partial f}{\partial y}$ at (10,10).

20. Find the stationary point for the function $f(x,y) = 5x^2 + 4y^2 + 2xy + x + 1$.

21. In congressional elections for 1974, the Republican percentage R of the Republican-Democratic vote in a district is given (approximately) by

$$R = 15.4725 + 2.5945\,E_R - 0.0804\,E_R^2 - 2.3648\,E_D + 0.0687\,E_D^2$$
$$+ 2.1914\,I_R - 0.0912\,I_R^2 - 0.8096\,I_D + 0.0081\,I_D^2 - 0.0277\,E_R I_R$$
$$+ 0.0493\,E_D I_D + 0.8579\,N - 0.0061\,N^2,$$

where E_R, E_D = campaign expenditures (in units of $10,000) by Republicans and Democrats, respectively

I_R, I_D = number of terms in Congress plus 1 for Republicans and Democrats, respectively

N = percentage of the two-party presidential vote that Nixon received in the district in 1968.

[From *Soc.Sci.Quart.*, Vol.58, No. 4 (1978), pp. 671-682]

In 1974, Congress set a limit of $188,000 for legal campaign expenditures.

(a) Based on this model, would you have advised a Republican candidate who had served 9 terms in Congress to spend $188,000 on his campaign? (Hint: look at $\frac{\partial R}{\partial E_R}$)

(b) What would be the expected percentage change for a Republican challenger (i.e. $I_R = 1$) who spent $150,000?

(c) Find the percentage above which the Nixon vote in a district had a negative effect on R. (Hint: look at when $\frac{\partial R}{\partial N} < 0$)

22. Find each of the following indefinite integrals:

(a) $\int x^3\,dx$ (b) $\int \sqrt{x}\,dx$ (c) $\int (x^2 - x + 3)\,dx$

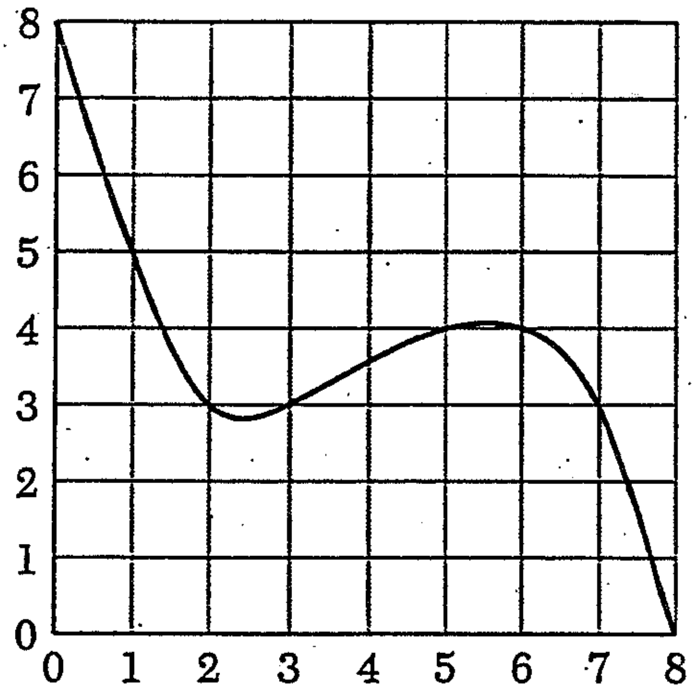
(d) $\int e^{2x}\,dx$ (e) $\int \frac{1}{x+1}\,dx$

23. Consider the function f which is graphed below. Which of the following is the best estimate of the definite integral

$$\int_1^6 f(x)\,dx \,?$$

(a) -24 (b) 9 (c) 20 (d) 38

Explain how you determined your answer.



24. Find the area under each of the following curves:

(a) $y = x^3 + 1$ from $x = -1$ to $x = 1$.

(b) $y = e^{-x}$ from $x = 0$ to $x = 1$.

25. (a) In 1990 the Lorenz curve in the United States for income in households was given approximately by $I(x) = x^{2.5}$. Calculate the Gini Index for that year.

(b) In 2000 the Gini Index for income was 0.460. What is a reasonable equation for the Lorenz curve?

26. One of the earliest pollution problems brought to the attention of the Environmental Protection Agency (EPA) involved a lake in South Dakota. For many years a small paper plant near the lake had been discharging waste containing carbon tetrachloride into the lake. At the time the EPA learned of the situation, the chemical was entering at a rate of 16 cubic yards per year. The agency ordered the installation of filters to slow (and eventually stop) the flow of carbon tetrachloride from the mill. The implementation of the program took three years, during which the flow of pollutant continued at a steady 16 cubic yards per year. Once the filters were installed the flow declined but from the time they were installed until the time the flow stopped, the rate of flow was approximated by

$$\text{Rate (in cubic yards per year)} = t^2 - 14t + 49,$$

where t is time measured in years since the EPA learned of the situation.

(a) Draw a graph showing the rate of carbon tetrachloride into the lake as a function of time, beginning at the time the EPA first learned of the situation.

(b) How many years elapsed between the time the EPA learned of the situation and the time the pollution flow stopped entirely?

(c) How much carbon tetrachloride entered the waters of the lake during the time shown in the graph in part (a)?

ANSWERS (CALCULUS)

1. (a) $y = 0.155t + 22.0$

   (b) The median age rose approximately 0.155 years each year beginning with 22.0 in 1980

   Note: The Census Bureau reports 25.3 for 2002 but the model would predict about 25.4

2. $p(t) = -0.8t + 57$ $\therefore p(20) = 41$. Note: The NHTS reported 41% for 2002.

3. (a) $\frac{476-117}{2002-1960} = \frac{359}{42} \approx 8.55$ (b) $\frac{96-117}{10} = -2.1$ (c) $\frac{111-111}{11} = 0$ (d) ≈ 11.88 (e) -0.4

4. (a) 1 (b) 0 (c) $\frac{f(2)-f(1)}{2-1} = \frac{3-4}{1} = -1$

5. $\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x)-f(x)}{\Delta x}$. So when $x=1$, $\frac{\Delta y}{\Delta x} = \frac{f(1+\Delta x)-f(1)}{\Delta x} = \frac{f(1+\Delta x)-4}{\Delta x}$

   (a)

Δx	3	2	1	0.1	0.01
$\frac{\Delta y}{\Delta x}$	1	0	-1	-1.9	-1.99

   (b) as Δx gets smaller, $\frac{\Delta y}{\Delta x}$ gets nearer to -2

6. (a) For $t=2$, $N(1)=6$ and $N(2)=40$ so $\frac{\Delta N}{\Delta t} = \frac{40-6}{1} = 34$

   also $N(2)=40$ and $N(3)=245$ so $\frac{\Delta N}{\Delta t} = \frac{245-40}{1} = 205$

   $\therefore N'(2)$ is between 34 and 205 and a reasonable estimate might be the average of the two values, i.e. $\frac{205+34}{2} = 120$. So $N'(2) \approx 120$

   Similarly, $N'(5) \approx 3632$ and $N'(8) \approx 120$.

   (b) In the first few days ($t=1,2,3$) the graph is almost flat; few people have heard the rumor so the number of new hearers is small, i.e. the change is small. Similarly, when $t=7,8,9$ the graph is almost flat; there is little change in the number of hearers since almost everyone in the town has heard the rumor.

   (c) Around day 5, the graph looks the steepest. Here about half the people have heard the rumor and can spread it to the half who have not heard it yet.

7. (a) 0 (b) $-\frac{3}{2}$ (c) $-8x$ (d) $y = x^{\frac{5}{3}}$ so $\frac{dy}{dx} = \frac{5}{3}x^{\frac{2}{3}}$

(e) $y = 8x^{-5}$ so $\frac{dy}{dx} = 8(-5x^{-6}) = -40x^{-6} = -\frac{40}{x^6}$

8. $f'(x) = 12x^3 - 12x = 12x(x^2-1) = 12x(x+1)(x-1)$

(a) $f'(2) = 24(3)(1) > 0$ so f is increasing at $x = 2$.

(b) $f'(-2) = 12(-2)(-1)(-3) = -72$ so the slope of the tangent line to the curve at $x = -2$ is -72. (Note: f is decreasing at $x = -2$)

(c) $f'(x) = 0$ when $x = 0$, $x = 1$, $x = -1$ so f is stationary at these values of x.

9. $f'(x) = 6x^2 - 6x - 12 = 6(x^2 - x - 2) = 6(x+1)(x-2)$

$f'(x) < 0$ for $-1 < x < 2$, i.e. f is decreasing if x is between -1 and 2.

10. (a) $\frac{dy}{dx} = (x^2-1)(3) + (3x+10)(2x) = 9x^2 + 20x - 3$

(b) $\frac{dy}{dx} = \frac{(5-2x^2)(2) - 2x(-4x)}{(5-2x^2)^2} = \frac{4x^2+10}{(5-2x^2)^2}$ (c) $6(2x-5)^2$

(d) $\frac{1}{2}(x^5-3)^{-\frac{1}{2}}(5x^4) = \frac{5x^4}{2\sqrt{x^5-3}}$

11. (a) at $x = \frac{3}{2}$ (b) $f''(x) = 12x^2 - 12 = 12(x^2-1) = 12(x+1)(x-1)$

$f''(-2) > 0$ and $f''(0) < 0$ so there is an inflection point at $x = -1$

$f''(0) < 0$ and $f''(2) > 0$ " " " " " " " $x = 1$

(c) at $x = 0$ and $x = \frac{3}{3}$ (d) $f''(x) = 12x^2$ so $x = 0$ is a possibility.

However, $f''(-1) > 0$ and $f''(1) > 0$ so there is no change in the concavity and hence no inflection point at $x = 0$.

12. (a) a local maximum at $x = 4$; $f(4) = 58$ is the maximum value.

(b) a local minimum at $x = 3$; $f(3) = -\frac{27}{2}$ is the minimum value.

also a local maximum at $x = -2$; $f(-2) = \frac{22}{3}$ is the maximum value.

(c) $f'(x) = x^3 - 16x = x(x^2-16) = x(x+4)(x-4)$; $f''(x) = 3x^2 - 16$

we test $x = 0$, $x = -4$, $x = 4$.

$f''(-4) > 0$ so there is a local minimum at $x = -4$; $f(-4) = -54$

$f''(4) > 0$ " " " " " " " $x = 4$; $f(4) = -54$

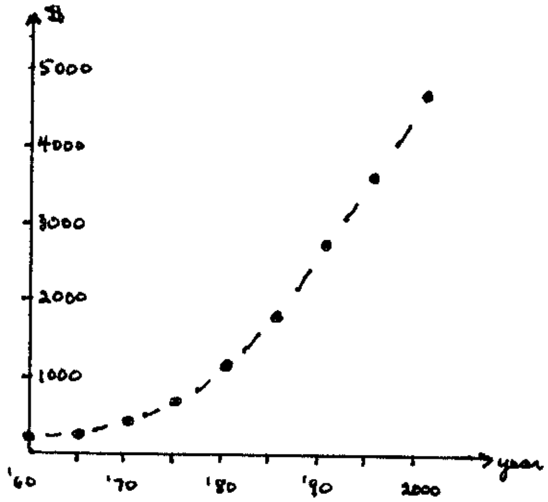
$f''(0) < 0$ so there is a local maximum at $x = 0$; $f(0) = 10$.

note: you can also test using the first derivative test.

13. (a) $P'(r) = -40{,}000r + 1{,}820{,}000$ $\therefore$ $P'(r) = 0 \Rightarrow r = \$45.50$

(b) $P(45.5) = \$39{,}655{,}000$

14.



(a) The graph looks more like a parabola (i.e. a quadratic function)

(b) $x = 40$ corresponds to 2000
$x = 42$ corresponds to 2002
Linear: $y(42) = 4211$
Quadratic: $y(42) = 5327$

note: HCFA projects \$5427 for 2002

15. (a) $P(60) = 16 + 84e^{-.66} \approx 59\%$ (b) $25 = 16 + 84e^{-0.011t} \Rightarrow e^{-0.011t} = \frac{9}{84} \Rightarrow t \approx 203.$

16. When $x = \frac{1}{2}$ there is a maximum of about -498.5

17. (a) $P(0) = P_0e^0 = P_0 = 3{,}050{,}000$ so $P_0 = 3{,}050{,}000$

$\therefore 6{,}874{,}000 = 3{,}050{,}000e^{20r}$ $\therefore r = \frac{1}{20}\ln\frac{6{,}874{,}000}{3{,}050{,}000} \approx 0.041$, a growth rate of about 4.1%

(b) $8{,}236{,}000 = 6{,}874{,}000e^{20r}$ $\therefore r \approx 0.009$, a growth rate of about 0.9%

18. $C'(x) = 0.01x - \frac{16}{x}$ $\therefore C'(x) = 0$ when $0.01x^2 = 16$ or $x = 40$

$C''(x) = 0.01 + \frac{16}{x^2}$ $\therefore C''(40) > 0$ so there is a local minimum when $x = 40$ and $C(40) \approx \$18.98$

19. (a) 0 (b) $\frac{f(11,10) - f(10,10)}{1} = \frac{8720 - 0}{1} = 8720$

(c) $\frac{\partial f}{\partial x} = 60x^2 + 20xy$. When $x = y = 10$ then $\frac{\partial f}{\partial x} = 8000$

(d) $\frac{f(10,11) - f(10,10)}{1} = \frac{-8930 - 0}{1} = -8930$

$\frac{\partial f}{\partial y} = -90y^2 + 10x^2$. When $x = y = 10$ then $\frac{\partial f}{\partial y} = -8000$

20. $\begin{cases} \frac{\partial f}{\partial x} = 10x + 2y + 1 \\ \frac{\partial f}{\partial y} = 8y + 2x \end{cases}$. Setting each partial derivative equal to 0 we get $\begin{cases} 10x + 2y = -1 \\ 2x + 8y = 0 \end{cases}$

Solving the simultaneous equations we obtain $x = -\frac{4}{38}$, $y = \frac{1}{38}$

$\therefore$ The stationary point is $\left(-\frac{4}{38}, \frac{1}{38}\right)$

21. (a) $\frac{\partial R}{\partial E_R} = 2.5945 - 0.1608 E_R - 0.0277\, I_R$

when $E_R = 18.8$ and $I_R = 10$ then $\frac{\partial R}{\partial E_R} < 0$

$\therefore$ R is decreasing; don't recommend.

(b) when $E_R = 15$ and $I_R = 1$ then $\frac{\partial R}{\partial E_R} = 0.1548$

(c) $\frac{\partial R}{\partial N} = 0.8579 - 0.0122 N$. If $\frac{\partial R}{\partial N} < 0$ then $N > 70.3$, i.e. about 70%

22. (a) $\frac{1}{4}x^4 + C$ (b) $\frac{2}{3}x^{\frac{3}{2}} + C$ (c) $\frac{1}{3}x^3 - \frac{1}{2}x^2 + 3x + C$

(d) $\frac{1}{2}e^{2x} + C$ (e) $\ln(x+1) + C$

23. Note that on the interval $[1,6]$ we have $2 < f(x) < 5$ so $2(5) < \int_1^6 f(x)\,dx < 5(5)$

So $\int_1^6 f(x)\,dx$ lies between 10 and 25 so the best estimate is (c).

alternately: count squares and portions of squares to get about 18.

24. (a) $A = \int_{-1}^{1} (x^3+1)\,dx = \frac{1}{4}x^4 + x \Big|_{-1}^{1} = \left(\frac{1}{4}+1\right) - \left(\frac{1}{4}-1\right) = 2$ square units.

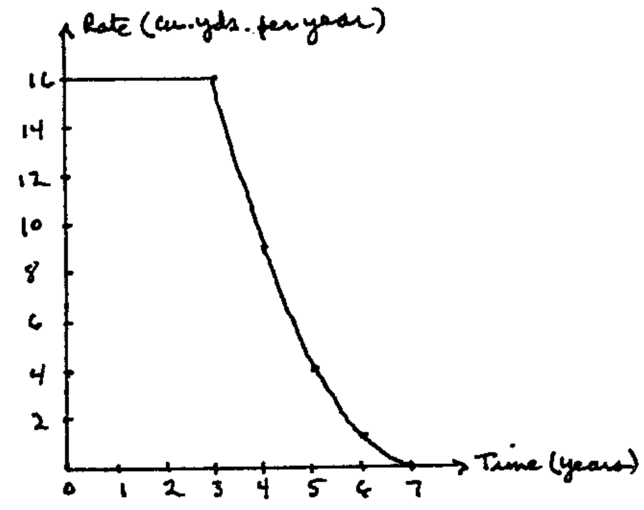
(b) $A = \int_0^1 e^{-x}\,dx = -e^{-x}\Big|_0^1 = -e^{-1} + 1 = 1 - \frac{1}{e}$ (this is about 0.63 square units)

25. (a) $GI = 1 - 2A = 1 - 2\int_0^1 x^{2.5}\,dx = 1 - 2\left(\frac{1}{3.5}\right)x^{3.5}\Big|_0^1 = 1 - 2\left(\frac{2}{7}\right) = \frac{3}{7} \approx 0.429$

(b) $0.460 = 1 - 2\int_0^1 x^k\,dx = 1 - \frac{2}{k+1}x^{k+1}\Big|_0^1 = 1 - \frac{2}{k+1}$ $\therefore \frac{2}{k+1} = 1 - 0.460 = 0.54$

$\therefore k \approx 2.7$ so $I(x) = x^{2.7}$

26. (a)



(b) 7 years

(c) $\int_0^3 16\,dt + \int_3^7 (t^2 - 14t + 49)t$

$= 16t\Big|_0^3 + \left(\frac{1}{3}t^3 - 7t^2 + 49t\right|_3^7$

$= 48 + 21\frac{1}{3}$

$= 69\frac{1}{3}$ cubic yards