1 An introduction to matrices

- An $r \times c$ matrix is a table of numbers, with r rows and c columns.
- For example
$$A = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 4 & -1 \end{pmatrix}$$
and
$$B = \begin{pmatrix} 5 & 1 \\ 0 & -2 \end{pmatrix}$$
are 3×2 and 2×2 matrices.
- The ijth entry in the matrix A, in the ith row and jth column, is denoted a_{ij}, so $a_{12} = 3$ and $b_{22} = -2$.
- A vector is a matrix with one column, a column of numbers
$$v = \begin{pmatrix} 6 \\ -1 \\ 4 \end{pmatrix}.$$
- Two matrices of the same size are added componentwise
$$\begin{pmatrix} 5 & 1 \\ 0 & -2 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} 8 & 1 \\ -2 & 2 \end{pmatrix}.$$
- If you multiply a matrix by a number, you multiply each entry in the matrix by the number
$$2A = 2\begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 4 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 6 \\ 4 & 0 \\ 8 & -2 \end{pmatrix}.$$
- Two matrices can be multiplied if the number of columns of the first matrix equals the number of rows of the second.

- If A is $r \times c$ and B is $c \times d$, then AB is $r \times d$.
- The ijth entry of the product AB is obtained by taking the product of the entries in the ith row of the first matrix with the entries in the j column of the second matrix, and adding them up.
- If C = AB, then

$$c_{ij} = \sum_{k=1}^{c} a_{ik} b_{kj}.$$

- With A 3×2 and B 2×2 as above, the product is 3×2

$$\begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 5 & -5 \\ 10 & 2 \\ 20 & 6 \end{pmatrix}.$$

- In general $BA \neq AB$, if fact in the example BA can't be done.
- The identity matrix I has 1's on the diagonal and zero's elsewhere

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- Multiplying by the identity matrix leaves the matrix unchanged

$$AI = IA = A.$$

- If it exists, the inverse of a square matrix M, denoted M^{-1} is such that

$$MM^{-1} = I = M^{-1}M.$$

- Matrix inverses are used to solve systems of linear equations, so if

$$Mx = b$$

  then

$$x = M^{-1}b$$

  provided the inverse exists.

- Inverses are hard to obtain in general.
- If M is diagonal (only the values on the diagonal are nonzero) the inverse is obtained by taking the reciprocal of each of the diagonal values
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & .5 & 0 \\ 0 & 0 & .333 \end{pmatrix}.$$
- If M is 2×2 there is a formula for its inverse
$$M^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$
- This will only work if $ad \neq bc$.
- For B above
$$B^{-1} = \begin{pmatrix} 5 & 1 \\ 0 & -2 \end{pmatrix}^{-1} = \frac{1}{-10} \begin{pmatrix} -2 & -1 \\ 0 & 5 \end{pmatrix}.$$
- A useful property of matrix inverses is that $(AB)^{-1} = B^{-1}A^{-1}$, provided the necessary inverses exist.
- The transpose of a matrix A, denoted A^T or A', is obtained by exchanging the rows and columns
$$\begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 4 & -1 \end{pmatrix}^{T} = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 0 & -1 \end{pmatrix}.$$
- A matrix which is equal to its transpose is called symmetric, for example
$$\begin{pmatrix} 5 & 1 \\ 1 & -2 \end{pmatrix}$$
is symmetric.

- The trace of a square matrix is the sum of its diagonal entries, so $tr(A) = \sum a_{ii}$.
- A useful properties of traces is that

$$tr(AB) = tr(BA)$$

  if the matrix products exist.

- An $r \times c$ matrix can be viewed as a function which maps vectors of length c into vectors of length r.
- If $\mathbf{x}$ is an $n \times k$ dimensional matrix with ij'th x_{ij}, then $\mathbf{x}$ looks like:

$$\mathbf{x} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ x_{31} & x_{32} & \dots & x_{3k} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}$$

2 Some matrix calculations in R

```
> cement=
+ read.csv(
+ "http://chase.mathstat.dal.ca/~bsmith/stat3340/Data/cement.c
+ header=T)
> cement
```

```
    i     y x_1 x_2 x_3 x_4
1   1  78.5   7  26   6  60
2   2  74.3   1  29  15  52
3   3 104.3  11  56   8  20
4   4  87.6  11  31   8  47
5   5  95.9   7  52   6  33
6   6 109.2  11  55   9  22
7   7 102.7   3  71  17   6
8   8  72.5   1  31  22  44
9   9  93.1   2  54  18  22
10 10 115.9  21  47   4  26
11 11  83.8   1  40  23  34
12 12 113.3  11  66   9  12
13 13 109.4  10  68   8  12
```

```
> y=cement[,2] #put the second column into a vector called y
> x=as.matrix(cement[,-c(1,2)]) #drop the first two columns
> x
```

```
      x_1 x_2 x_3 x_4
 [1,]   7  26   6  60
 [2,]   1  29  15  52
 [3,]  11  56   8  20
 [4,]  11  31   8  47
 [5,]   7  52   6  33
```

```
 [6,]  11  55   9  22
 [7,]   3  71  17   6
 [8,]   1  31  22  44
 [9,]   2  54  18  22
[10,]  21  47   4  26
[11,]   1  40  23  34
[12,]  11  66   9  12
[13,]  10  68   8  12
```

```
> dim(x) #dimension of x
```

```
[1] 13  4
```

```
> ones=rep(1,length(y))
> ones
```

```
 [1] 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
> dim(ones)
```

```
NULL
```

```
> X=cbind(ones,x)  #cbind concatenates matrices by columns
> X
```

```
      ones x_1 x_2 x_3 x_4
 [1,]    1   7  26   6  60
 [2,]    1   1  29  15  52
 [3,]    1  11  56   8  20
 [4,]    1  11  31   8  47
 [5,]    1   7  52   6  33
 [6,]    1  11  55   9  22
 [7,]    1   3  71  17   6
 [8,]    1   1  31  22  44
 [9,]    1   2  54  18  22
```

```
[10,]    1  21  47   4  26
[11,]    1   1  40  23  34
[12,]    1  11  66   9  12
[13,]    1  10  68   8  12
```

```
> dim(X)
```

```
[1] 13  5
```

The entry in row i and column j of X is the observation on the $j-1$'st predictor variable on subject i, where we think of the column of ones as the 0'th predictor variable.

```
> xp=t(X) #X' = transpose of X
> xp
```

```
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,
ones    1    1    1    1    1    1    1    1    1     1     1
x_1     7    1   11   11    7   11    3    1    2    21     1
x_2    26   29   56   31   52   55   71   31   54    47    40
x_3     6   15    8    8    6    9   17   22   18     4    23
x_4    60   52   20   47   33   22    6   44   22    26    34
```

```
> dim(xp)
```

```
[1]  5 13
```

```
> xpx=xp%*%X #multiply transpose of X by X
> xpx
```

```
     ones  x_1   x_2  x_3   x_4
ones   13   97   626  153   390
x_1    97 1139  4922  769  2620
x_2   626 4922 33050 7201 15739
x_3   153  769  7201 2293  4628
x_4   390 2620 15739 4628 15062
```

```
> xpxi=solve(xpx) #find the inverse of X'X
> round(xpxi,3)

        ones    x_1    x_2    x_3    x_4
ones 820.655 -8.442 -8.458 -8.635 -8.290
x_1   -8.442  0.093  0.086  0.093  0.084
x_2   -8.458  0.086  0.088  0.088  0.086
x_3   -8.635  0.093  0.088  0.095  0.086
x_4   -8.290  0.084  0.086  0.086  0.084

> round(xpxi%*%xpx,3) # (X'X)inverse times (X'X) should be the

     ones x_1 x_2 x_3 x_4
ones    1   0   0   0   0
x_1     0   1   0   0   0
x_2     0   0   1   0   0
x_3     0   0   0   1   0
x_4     0   0   0   0   1

> b=xpxi%*%t(X)%*%y  #(X'X)inverse X'y
> b

           [,1]
ones 62.4053693
x_1   1.5511026
x_2   0.5101676
x_3   0.1019094
x_4  -0.1440610

> print(lm(y~x_1+x_2+x_3+x_4,data=cement)) #least squares esti

Call:
lm(formula = y ~ x_1 + x_2 + x_3 + x_4, data = cement)
```

```
Coefficients:
(Intercept)          x_1          x_2          x_3          x_4
    62.4054       1.5511       0.5102       0.1019      -0.1441

>
```