Statistics 5041
5. Inverses and Determinants

Gary W. Oehlert
School of Statistics
313B Ford Hall
612-625-1557
gary@stat.umn.edu

A few (final?) ideas about matrices.

A square matrix is *symmetric* if $\mathbf{X} = \mathbf{X}'$. In that case, the first row equals the first column, the second row equals the second column, and so on.

$$\mathbf{X} = \begin{bmatrix} 9.8 & 12.6 & 7.0 & 8.9 \\ 12.6 & 15.6 & 10.1 & 8.9 \\ 7.0 & 10.1 & 11.2 & 8.0 \\ 8.9 & 8.9 & 8.0 & 11.6 \end{bmatrix}$$

If $\mathbf{X}$ is square, $\mathbf{X} + \mathbf{X}'$ is always symmetric.

If $\mathbf{X}$ is symmetric, $\mathbf{X} = .5(\mathbf{X} + \mathbf{X}')$.

A *diagonal* matrix is a square matrix $\mathbf{X}$ with $X_{ij} = 0$ whenever $i \neq j$. That is, all the off-diagonal elements are zero.

$$\mathbf{X} = \begin{bmatrix} 9.8 & 0 & 0 & 0 \\ 0 & 15.6 & 0 & 0 \\ 0 & 0 & 11.2 & 0 \\ 0 & 0 & 0 & 11.6 \end{bmatrix}$$

If v is a vector of length n, then $\text{diag}(v)$ is the $n \times n$ diagonal matrix with the elements of v down the diagonal. In MacAnova,

```
Cmd> dmat(vector(9.8,15.6,11.2,11.6))
(1,1)       9.8          0          0          0
(2,1)         0       15.6          0          0
(3,1)         0          0       11.2          0
(4,1)         0          0          0       11.6
```

The *identity matrix* is a diagonal matrix with ones down the diagonal.

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

The identity matrix is to matrix multiplication what 1 is to regular multiplication.

Let $\mathbf{X}$ by $m \times n$, let $\mathbf{Y}$ be $n \times p$, and let $\mathbf{Z}$ be $n \times n$. Then

$$\mathbf{X}\mathbf{I}_n = \mathbf{X}$$

and

$$\mathbf{I}_n\mathbf{Y} = \mathbf{Y}$$

and

$$\mathbf{I}_n\mathbf{Z} = \mathbf{Z} = \mathbf{Z}\mathbf{I}_n.$$

You can multiply on the left or right, as long as dimensions match.
In MacAnova

```
Cmd> I4 <- dmat(vector(1,1,1,1));I4
(1,1)          1         0         0         0
(2,1)          0         1         0         0
(3,1)          0         0         1         0
(4,1)          0         0         0         1

Cmd> I4 <- dmat(rep(1,4)) # same thing

Cmd> I4 <- dmat(4,1) # same thing

Cmd> X
(1,1)       11.4      11.2      10.2      10.8
(2,1)       11.6        11      12.4       9.4

Cmd> X %*% I4
(1,1)       11.4      11.2      10.2      10.8
(2,1)       11.6        11      12.4       9.4

Cmd> Y
(1,1)       12.6      10.6      10.4
(2,1)         12       8.6       6.4
(3,1)       12.6      12.6      10.4
(4,1)         10        11        12

Cmd> I4 %*% Y
(1,1)       12.6      10.6      10.4
(2,1)         12       8.6       6.4
(3,1)       12.6      12.6      10.4
(4,1)         10        11        12
```

Let $\mathbf{1}_n$ be a column vector of n ones.

$$\mathbf{1}_n = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

If x is a vector of length n, then

$$< \mathbf{1}_n, x >= \mathbf{1}_n' x = \sum_{i=1}^{n} 1 \times x_i = \sum_{i=1}^{n} x_i$$

expresses the sum of the elements of x in matrix form.

```
Cmd> x <- vector(2.3, 8.7, 5.4, 6.9)

Cmd> one4 <- rep(1,4)

Cmd> one4' %*% x
(1,1)         23.3

Cmd> sum(x)
(1)         23.3
```

Vectors $v_1, v_2, \ldots, v_k$ (all $n \times 1$) are *linearly independent* if none of them can be written as a linear combination of the others. That is, we cannot write

$$\begin{aligned} v_j &= c_1 v_1 + c_2 v_2 + \ldots + c_{j-1} v_{j-1} + \\ &\quad c_{j+1} v_{j+1} + c_{j+2} v_{j+2} + \ldots + c_k v_k \end{aligned}$$

for any $j, 1 \leq j \leq k$ and any set of coefficients c_i.
Equivalently, $0 = \sum_{i=1}^{k} c_i v_i$ implies that all the c_i's are zero.
Let $\mathbf{X}$ be $n \times p$. The *rank* of $\mathbf{X}$ can be defined in many ways, including:

The maximum number of linearly independent columns of $\mathbf{X}$.

The maximum number of linearly independent rows of $\mathbf{X}$.

The minimum number of outer products (of vectors) needed to reconstruct $\mathbf{X}$ when summed.

$$\text{rank}(\mathbf{X}) \leq \min(n, p)$$

$\mathbf{X}$ has *full rank* if rank$(\mathbf{X}) = \min(n, p)$.
Some special cases:

- A nonzero row or column vector has rank 1.
- If vectors v and w are nonzero, then vw' has rank 1.
- If vectors $\ell_1, \ell_2, \ldots, \ell_k$ are linearly independent, and vectors $r_1, r_2, \ldots, r_k$ are linearly independent, then

  $$\mathbf{X} = \sum_{i=1}^{k} \ell_i r_i'$$

  has rank k.

Let $\mathbf{A}$ be $n \times n$. Suppose that there exists an $n \times n$ matrix $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$$

Then $\mathbf{B}$ is the *inverse* of $\mathbf{A}$, written

$$\mathbf{A}^{-1}$$

$\mathbf{A}^{-1}$ is unique.
$\mathbf{A}^{-1}$ exists if and only if $\mathbf{A}$ has full rank.

```
Cmd> A
(1,1)        6        2        8
(2,1)        8        5        6
(3,1)        1        2        3

Cmd> B<-solve(A);B # inversion
(1,1)  0.042857   0.14286      -0.4
(2,1)  -0.25714   0.14286       0.4
(3,1)   0.15714  -0.14286       0.2

Cmd> A%*%B
(1,1)         1         0   3.3e-16
(2,1)         0         1   4.4e-16
(3,1)   2.7e-17         0         1

Cmd> B%*%A
(1,1)         1         0  -5.5e-17
(2,1)   1.1e-16         1   2.2e-16
(3,1)   5.5e-17   5.5e-17         1

Cmd> c1 <- vector(3,5,1);\
c2 <- vector(1,1,6)

Cmd> A <- hconcat(c1,c2,c1+c2); solve(A)
(1,1)   2.3e+15  -1.3e+15  -1.5e+14
(2,1)   2.3e+15  -1.3e+15  -1.5e+14
(3,1)  -2.3e+15   1.3e+15   1.5e+14

Cmd> A %*% solve(A)
(1,1)        0        0        0
(2,1)        0        0        0
(3,1)        0        0        0
```

Oops. Nonsense result because numerical singularity not found.
For every square matrix $\mathbf{X}$ ($k \times k$) we can compute a quantity called the *determinant* and denoted by $|\mathbf{X}|$.

$$\det(\mathbf{X}) = \sum_{p \in P} (-1)^{s(p)} X_{1p_1} X_{2p_2} \cdots X_{kp_k}$$

where p is a permutation of the numbers 1 through k, P is the set of all such permuations, and $s(p)$ is the number of permuation inversions in the permutation p.

$$s(p) = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} I(p_i > p_j)$$

Not very intuitive.

$$|\mathbf{X}| = \left| \begin{bmatrix} 1 & 5 \\ 3 & 10 \end{bmatrix} \right| = 1 \times 10 - 3 \times 5 = -5$$

Permuations are (1,2) and (2,1), with 0 and 1 inversions.
For a 3×3, the permuations are

p	$s(p)$
(1,2,3)	0
(1,3,2)	1
(2,1,3)	1
(2,3,1)	2
(3,1,2)	2
(3,2,1)	3

$$\begin{aligned} |\mathbf{X}| &= \left| \begin{bmatrix} 1 & 5 & 3 \\ 3 & 10 & 4 \\ 5 & 8 & 5 \end{bmatrix} \right| \\ &= 1 \times 10 \times 5 \\ &\quad -1 \times 4 \times 8 \\ &\quad -5 \times 3 \times 5 \\ &\quad +5 \times 4 \times 5 \\ &\quad +3 \times 3 \times 8 \\ &\quad -3 \times 10 \times 5 \\ &= -35 \end{aligned}$$

The definitional form for the determinant is not computationally efficient. Different methods are used in software.

```
Cmd> X
(1,1)          1          3          5
(2,1)          5         10          8
(3,1)          3          4          5

Cmd> det(X)
(1)            -35
```

Facts about determinants.
$\mathbf{X}$ (square) has an inverse if and only if $\det(\mathbf{X}) \neq 0$.
If $\mathbf{X}^{-1}$ exists, $\det(\mathbf{X}^{-1}) = 1/\det(\mathbf{X})$.
Adding a multiple of a row (or column) to any other row (or column) does not change the determinant.
Multiplying a row by a scalar multiplies the determinant by that scalar.

```
Cmd> det(hconcat(c1,c2,c1+c2))
(1)  -1.2603e-14
```

```
Cmd> det(solve(X))
(1)    -0.028571

Cmd> 1/det(X)
(1)    -0.028571

Cmd> X2 <- X;X2[1,] <- 2*X2[1,];det(X2)
(1)          -70

Cmd> X2 <- X;X2[1,] <- -2*X2[1,];det(X2)
(1)           70

Cmd> X3 <- X;X3[,1] <- X3[,1] + 2*X3[,3];det(X3)
(1)          -35
```

The *trace* of a square matrix is the sum of the diagonal elements.

$$\text{trace}(\mathbf{X}) = \sum_{i=1}^{k} X_{ii}$$

```
Cmd> trace(X)
(1)           16
```