Winter 2011

Foote

6. Linear Algebra I

I. Vector and Matrix Basics

A. Matrix ($n \times p$) as set of n row vectors in p-space and p column vectors in n-space.

1. By convention, let's focus on column vectors.
   - Then the *vector space* $\mathbf{R}^n$ consists of *all* column vectors with n elements.
     Example

$$\mathbf{X} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$

$$\vec{x_1} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \vec{x_2} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \vec{x_3} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

   - The rows can be thought of as the columns of the transpose of the matrix, $\mathbf{X}^T$, where $x_{ij}^T = x_{ji}$
2. Operations on vectors that result in vectors still in the vector space:
   a. Addition of any two (or more) vectors
      •$\vec{a} = \vec{x} + \vec{y}$ where $a_i = x_i + y_i$
      •Implementing in R: `a<-x+y`
   b. Multiplication of vectors by scalars
      •$\vec{a} = c\vec{x}$ where $a_i = cx_i$
      •Implementing in R: `a<-c*x`
3. Subspace of vector space defined by two conditions:
   a. If $\vec{x}$ and $\vec{y}$ are two vectors in the subspace, then $\vec{x} + \vec{y}$ is in the subspace.

and

   b. If $\vec{x}$ is in the subspace and c is a scalar, then $c\vec{x}$ is in the subspace.
4. Clearly (3a) and (3b) imply that any linear combination of vectors in a subspace results in a vector that is in that subspace.
5. Example of subspace: x-y plane within x-y-z space.

B. Linear Dependence and Independence

1. Take a set of k vectors $\vec{x}_1, ..., \vec{x}_k$ and a set of k scalars $c_1, ..., c_k$.
2. If the only way to get the sum $c_1\vec{x}_1 + ... + c_k\vec{x}_k = 0$ is to set all the scalars $c_1 = c_2 = ... = c_k = 0$, then the vectors are linearly independent.
3. In other words, if one or more of the vectors can be formed as a (nontrivial) linear combination of (some of) the others, then the vectors are linearly dependent.
4. Note difference between this concept of independence and that in statistics, in which independence implies no discernible correlation, i.e., no clear evidence for angle between vectors differing from $\pi/2$.

C. Rank and Related Concepts

1. The number of linearly independent columns is the dimension or rank r of the column space of a matrix.
   - $r \leq \min(n, p)$ for an $n \times p$ matrix.
   - Examples

$$\begin{pmatrix} 1 & 2 & 3 \\ 7 & 6 & 4 \end{pmatrix} \text{rank=2}$$

$$\begin{pmatrix} 2 & 7 & 3 \\ 4 & 6 & 3 \\ 9 & 1 & 1 \\ 8 & 1 & 1 \end{pmatrix} \text{rank=3}$$

$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 3 & 6 & 1 \\ 4 & 8 & 1 \end{pmatrix}$ rank=2, because column 2 is a linear multiple of column 1.

2. The rank of the column space is exactly the same as the rank of the row space. I.e., the number of linearly independent rows is exactly equal to the number of linearly independent columns.
3. Spanning set: Consider the set of vectors $\vec{x}_1, ...\vec{x}_k$. If a vector space V consists of all possible linear combinations of the $\vec{x}$'s, then the $\vec{x}$'s *span* the vector space V.
4. Basis: A set of vectors that is linearly independent and spans a vector space forms a *basis* for that space.
   a. Illustration with 1, 2, and 3 vectors in x-y plane. Let $\vec{v}_1, \vec{v}_2$, and $\vec{v}_3$ be the vectors $\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, and $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.
   •Any one of these vectors is linearly independent but does not span the space.
   •The three vectors together span the space, but they are not independent.
   •Any pair of these vectors spans the space and is linearly independent, and thus forms a basis for the space.
   b. Spanning sets and bases are not unique. Special consideration often given to *orthonormal* bases, in which basis vectors are orthogonal and have unit length: e.g. $\vec{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ for the x-y plane.

D. Additional Vector Operations

- See above for addition and multiplication by scalar

1. Inner product (dot product, minor product): result is scalar, $a = \vec{x} \cdot \vec{y} = \vec{x}^T \vec{y} = \sum_i x_i y_i$
   - Implementing in R: `a<-t(x)%*%y` or `a<-crossprod(x,y)` or `a<-sum(x*y)`
   - Length of $\vec{x}$, denoted $\|\vec{x}\|$, is equal to $\sqrt{\vec{x} \cdot \vec{x}}$.
2. Outer product: result is a matrix, $\mathbf{A} = \vec{x}\vec{y}^T$, where $a_{ij} = x_i y_j$
   - Implementing in R: `a<-x%*%t(y)` or `a<-x%o%y` for vectors `x` and `y`. Note that for matrices `X` and `Y`, `X%o%Y` is not the same as `X%*%t(Y)`.

E. Matrix manipulation

1. Addition: $\mathbf{C} = \mathbf{A} + \mathbf{B}$, where $c_{ij} = a_{ij} + b_{ij}$
   - Implementing in R: `C<-A+B`
   - Associative
   - Commutative
   - Addition with scalar: $\mathbf{C} = \mathbf{A} + b$, where $c_{ij} = a_{ij} + b$
   •Implementing in R: `C<-A+b`
2. Multiplication: $\mathbf{C} = \mathbf{AB}$, where $c_{ij} = \sum_{k=1}^{p_A} a_{ik} b_{kj}$
   Example

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 1 & 1 \end{pmatrix} \mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\mathbf{AB} = \begin{pmatrix} (1 \cdot 1 + 0 \cdot 1) & (1 \cdot 2 + 0 \cdot 1) & (1 \cdot 3 + 0 \cdot 1) \\ (2 \cdot 1 + 1 \cdot 1) & (2 \cdot 2 + 1 \cdot 1) & (2 \cdot 3 + 1 \cdot 1) \\ (1 \cdot 1 + 1 \cdot 1) & (1 \cdot 2 + 1 \cdot 1) & (1 \cdot 3 + 1 \cdot 1) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 5 & 7 \\ 2 & 3 & 4 \end{pmatrix}$$

   - Defined only if $p_A = n_B$
   - Implementing in R: `C<-A%*%B`.
   - Generally not commutative, except for multiplication by **I** and multiplication of two diagonal matrices
   - Associative, if number of rows and columns allows the operation

- Distributive
- Multiplication by scalar: $\mathbf{A} = c\mathbf{X}$, where $a_{ij} = cx_{ij}$
  - Implementing in R: `A<-c*X`

3. Kronecker Product (direct product, tensor product):

$$\mathbf{A}_{(\mathbf{n}\times\mathbf{m})} \otimes \mathbf{B}_{(\mathbf{p}\times\mathbf{q})} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1m}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2m}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \cdots & a_{nm}\mathbf{B} \end{pmatrix},$$

where the resulting product has dimensions $(np \times mq)$ and the term $a_{ij}\mathbf{B}$ is a block of the matrix equal to:

$$\begin{pmatrix} a_{ij}b_{11} & a_{ij}b_{12} & \cdots & a_{ij}b_{1q} \\ a_{ij}b_{21} & a_{ij}b_{22} & \cdots & a_{ij}b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{ij}b_{p1} & a_{ij}b_{p2} & \cdots & a_{ij}b_{pq} \end{pmatrix}$$

- Implementing in R: `A%x%B`.

4. Hadamard Product (Schur Product): $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$, where $c_{ij} = a_{ij}b_{ij}$. I.e., it is elementwise multiplication.
   a. Note the symbol $\cdot$ to indicate the Hadamard product.
   b. Operation defined if $n_A = n_B$ and $p_A = p_B$.
   c. This is what is done in R if you simply type `A*B`.

F. Some Special Matrices

1. Square: $n = p$
2. Symmetrical: square; $\mathbf{X}^T$=$\mathbf{X}$
3. Diagonal: symmetrical; $x_{ij} = 0$ for $i \neq j$
4. Identity: diagonal; $\mathbf{I}$ such that $I_{ii} = 1$
5. Inverse: $\mathbf{X}^{-1}$ such that $\mathbf{X}\mathbf{X}^{-1}$=$\mathbf{X}^{-1}\mathbf{X}$=$\mathbf{I}$
6. Orthogonal: square; all columns orthonormal; all rows orthonormal
   - $\mathbf{Q}\mathbf{Q}^T$=$\mathbf{I}$; $\mathbf{Q}^T\mathbf{Q}$=$\mathbf{I}$; $\mathbf{Q}^T$=$\mathbf{Q}^{-1}$
   - Example: rotation matrix (counterclockwise by angle θ)

$$\mathbf{Q} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}; Q^T = Q^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

•For example, let $\theta = \pi/4$, so that $\cos\theta = \sin\theta = \sqrt{0.5}$, and let $\mathbf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$. Then:

$$\mathbf{QX} = \begin{pmatrix} 0.707 & -0.707 \\ 0.707 & 0.707 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -0.707 & -0.707 \\ 0.707 & 2.121 \end{pmatrix}$$

.

- Example: permutation matrix, which produces a row exchange. For example,

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}; \mathbf{P}^T = \mathbf{P}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

•The column j in which the 1 falls for a particular row i tells which row will take the place of row i in the permuted matrix. For example, in $\mathbf{P}$ above, row 2 goes to row 1, row 3 goes to row 2, and row 1 goes to row 3. Thus:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 3 \end{pmatrix}$$