**BIOS 667: Longitudinal Data Analysis**
Vectors and Matrices

Matrix notation is used in this class because it provides compact expressions that are easy to read and understand. The approach here is a practical one with very minimal theory involved.

We will only deal with vectors and matrices of real-values. The term *scalar* means one number such as 2.7.

All vectors will be column vectors. A set of real numbers such as $a_1, \ldots, a_n$ can be collected into a vector denoted $a = (a_1, \ldots, a_n)^\top$. This vector has length $n$. The superscript $^\top$ denotes the transpose. The notation $(a_1, \ldots, a_n)$ represents a row vector, while $(a_1, \ldots, a_n)^\top$ is a column vector.

A special vector is a vector of 1's, denoted $1_n$ or $1_{n\times1}$ or just 1 provided there is no confusion with the scalar 1.

Another special vector is the zero vector, a vector of zeros, $0_{n\times1}$.

Vectors of the same length can be added and subtracted elementwise, $a = b + c, d = b - c$ mean that $a_i = b_i + c_i, d_i = b_i - c_i$, and $a, b, c$ and $d$ have the same length.

The *inner product* is an operation on two vectors of the same length $n$, defined as $a^\top b = b^\top a = \sum_{i=1}^n a_i b_i$. The result is a scalar (one number).

A matrix is a rectangular array of numbers. In matrix $A$, the entry in the $i$th row and $j$th column is $a_{ij}$.

A matrix $A$ with 10 rows and 3 columns is a $10 \times 3$ matrix. The pair of numbers $(10, 3)$ is called the *dimension* or *dimensions* of the matrix. The transpose of $A$ is $A^\top$ which is a $3 \times 10$ matrix with $a_{ji}$ in the $i$th row and $j$th column; $i$ goes from 1 to 3 while $j$ goes from 1 to 10.

If the number of rows equals the number of columns the matrix is called a *square* matrix.

If $A = A^\top$, then $A$ is said to be *symmetric*. Covariance matrices are symmetric.

Matrices of the same dimension can be added and subtracted elementwise, $A = B + C, D = B - C$ mean that $a_{ij} = b_{ij} + c_{ij}, d_{ij} = b_{ij} - c_{ij}$, and $A, B, C$ and $D$ must have the same dimension.

The product of matrices $A$ and $B$ is denoted $AB$ and is defined only if the number of columns in $A$ is equal to the number of rows in $B$. If $A$ is $m \times n$ and $B$ is $n \times p$, then the product $C = AB$ is an $m \times p$ matrix ($m$ rows and $p$ columns), and element $c_{ij}$ in $C$ is the inner product of the $i$th row of $A$ and the $j$th column of $B$.

Order of matrix multiplication does not matter, $ABC = A(BC) = (AB)C$. Matrix multiplication is not commutative, generally $AB \neq BA$.

Note that if $a_{m\times1}$ and $b_{n\times1}$ are vectors then $ab^\top$ is an $m \times n$ matrix.

A special square matrix is the identity matrix, usually denoted $I_{n\times n}$, $I_n$ or just $I$ if there is no confusion about its dimension.

Another special matrix is a matrix of 1's, usually denoted $J_{m\times n}$ and it can be represented by $1_m 1_n^\top$. Since $J$ need not be square the notation $J_n$ should be avoided.

If there is no linear dependence among the columns (or rows) of an $n \times n$ matrix $A$, we say that $A$ is nonsingular.

If $A$ and $B$ are $n \times n$ matrices and $AB = BA = I$, we say that $B$ is the inverse of $A$ and express that by $B = A^{-1}$. The inverse exists only if $A$ is nonsingular. The inverse is unique (when it exists).

Exercises:
Let $a = (1, 2, 3, 4, 5, 6)^\top$, $b = (2, -1, 4, -3, 6, -5)^\top$.
Compute $a^\top a, b^\top b, a^\top b, a1_4^\top, 1^\top a, 1^\top(aa^\top)1, J_{4\times 6}a$.

What is the inverse of $A^{-1}$?

Random vectors and matrices are vectors and matrices whose elements are random variables. For example, $Y = (Y_1, Y_2, Y_3, Y_4)^\top$ is a random vector. Its mean is $\mathrm{E}[Y] = (\mu_1, \mu_2, \mu_3, \mu_4)^\top$. With multivariate observations $Y_1, \ldots, Y_n$, each $Y_i$ is itself a vector. In the TLC study each $Y_i$ is a $4 \times 1$ vector. The sample mean vector, as a random vector, is $\bar{Y}$, and its realized (observed) value is $\bar{y}$ The sample covariance matrix $S$ is a $4 \times 4$ matrix, and its realized value is $s$. Note that in the univariate case the common notation is different, $S^2$ and $s^2$.

If the random vector $Y_{n\times 1}$ has mean vector $\mu_{n\times 1} = \mathrm{E}[Y]$ and covariance matrix $\Sigma_{n\times n} = \mathrm{cov}(Y)$, and $A_{m\times n}$ is a matrix, then $\mathrm{E}[AY] = A\mu$ and $\mathrm{cov}(AY) = A\Sigma A^\top$ It follows that if $a$ is an $n \times 1$ vector then the linear combination $a^\top Y$ is a scalar random variable with mean $\mathrm{E}[a^\top Y] = a^\top \mu$ and variance $\mathrm{var}(a^\top Y) = a^\top \Sigma a$.

Expressions of the type $a^\top Ba$ where $B$ is a matrix are called *quadratic forms*. Examples include sums of squares in ANOVA and regression and many test statistics.

A square matrix $A_{n\times n}$ is said to be *positive definite* if $x^\top Ax > 0$ for all non-zero vectors $x$ (of length $n$). If $A$ is the covariance matrix of a random vector $Y$, positive definiteness of $A$ guarantees that all non-zero linear combinations $a^\top Y$ have strictly positive variance, i.e. $\mathrm{var}(a^\top Y)$ can't be negative and can't be zero unless $a = 0$.

A square matrix $A_{n\times n}$ is said to be *positive semidefinite* if $x^\top Ax \geq 0$ for all non-zero vectors $x$.

If $\Sigma$ is a positive definite $n \times n$ covariance matrix, there exists a unique lower-triangular matrix $L$ with positive diagonal elements such that $\Sigma = LL^\top$. Matrix $L$ is known as the Choleski root of $\Sigma$. The Cholesky root of $A$ can be computed by "chol(A)" in both SAS/IML and R.