Lecture 2
Linear Algebra

Symmetric matrices A $(n \times n)$

One obtains the n eigenvalues of an $n \times n$ matrix A by solving the characteristic equation $|A - \lambda I| = 0$ for λ, where λ is a complex scalar. One obtains the n eigenvectors b by solving $Ab = \lambda b$.

Lemma 1: For real symmetric matrices, the eigenvalues are real and we can choose the eigenvectors to be orthonormal.

Proof: First we show that the eigenvectors are orthogonal. We have $Ab_1 = \lambda_1 b_1$ and $Ab_2 = \lambda_2 b_2$. Hence, $\lambda_1 b_1' A b_2 = b_1' A' \lambda_2 b_2 \Rightarrow \lambda_1 b_1' A b_2 = \lambda_2 b_1' A b_2 \Rightarrow \lambda_1 = \lambda_2$ or $b_1' A b_2 = 0$. Suppose $\lambda_1 \neq \lambda_2$, then $0 = b_1' A b_2 = b_1' \lambda_2 b_2 = \lambda_2 b_1' b_2$ and $0 = \lambda_1 b_1' b_2$. By an analogous argument, $0 = \lambda_2 b_1' b_2$. Hence, $b_1' b_2 = 0$. Alternatively, suppose $\lambda_1 = \lambda_2$. Then, $\alpha b_1 + \gamma b_2$ solves $Ab = \lambda b$ for all $\alpha, \gamma \in R$. That is, there exists a vector space of solutions. We can take an orthogonal basis of this vector space, i.e., take vectors $b_1, b_2, ..., b_m$, where m is the multiplicity of roots such that $\lambda_1 = \lambda_2 = \cdots = \lambda_m$. Finally, if we take all the solutions to $Ab_i = \lambda_i b_i$ and normalize them to have length one, then $b_1, ..., b_n$ are orthonormal and $B = [b_1, ..., b_n]$ is an orthogonal matrix.

Lemma 2: Any symmetric real matrix A can be diagonalized using an orthogonal matrix. That is, there exists B orthogonal and Λ diagonal such that $B'AB = \Lambda$.

Proof: Take the matrix B from above. Let $\Lambda = \text{diagonal}(\lambda_1, ..., \lambda_n)$. Then, $AB = [Ab_1, Ab_2, ..., Ab_n] = [b_1\lambda_1, b_2\lambda_2, ..., b_n\lambda_n] = B\Lambda$. Pre-multiply by B' to obtain $B'AB = B'B\Lambda = \Lambda$. $\square$

Note that for an arbitrary (i.e., non-symmetric $r \times s$) matrix A, there exists a singular value decompositon: $BAC = \Lambda$, where B and C are orthogonal and $\Lambda = [I \vdots 0]$ is "diagonal."

Lemma 3: For A symmetric, $|A| = \pi_{i=1}^{n} \lambda_i$. That is, the determinant of a symmetric matrix equals the product of its eigenvalues.

Proof: Write A as $A = B\Lambda B'$, then $|A| = |B\Lambda B'| = |B| \cdot |\Lambda| \cdot |B'| = |\Lambda|$ since the determinant of an orthogonal matrix is plus or minus one. This follows because $B'B = I \Rightarrow |B| \cdot |B'| = |I| = 1 \Rightarrow |B|^2 = 1 \Rightarrow |B| \pm 1$. $\square$

Lemma 4: For A symmetric, $\text{tr}(A) = \sum_{i=1}^{n} \lambda_i$. That is, the trace of a symmetric matrix equals the sum of its eigenvalues.

Proof: $\text{tr}(A) = \text{tr}(B\Lambda B') = \text{tr}(\Lambda B'B) = \text{tr}(\Lambda) = \sum_{i=1}^{n} \lambda_i$. $\square$

Positive Definite Matrices

A matrix A is positive definite if $x'Ax > 0 \ \forall x \neq 0$. A matrix A is positive semi-definite if $x'Ax \geq 0 \ \forall x \neq 0$. Clearly, if A is positive definite, it is nonsingular (otherwise $Ax = 0$ for some x).

Lemma 5: If A is positive definite (positive semi-definite), then all its eigenvalues are positive (non-negative).

Proof: If $b_1, ..., b_n$ are the eigenvectors of A corresponding to $\lambda_1, ..., \lambda_n$, then $0 < b_j'Ab_j = \lambda_j b_j' b_j = \lambda_j$. $\square$

Lemma 6: If A is symmetric and positive semi-definite, then there exists an $n \times n$ matrix C such that $A = CC'$.

Proof: $A = B\Lambda B' = B\Lambda^{1/2}\Lambda^{1/2}B' = CC'$, where $\Lambda^{1/2} = \text{diagonal}(\lambda_1^{1/2}, ..., \lambda_n^{1/2})$ and $C = B\Lambda^{1/2}$. This only works if $\lambda_j \geq 0 \ \forall j$. $\square$

Note that if A is symmetric and positive semi-definite, then A^r is defined for all $r \in R$: $A^r = B\Lambda^r B'$.

Lemma 7: If A is symmetric and positive definite, then so is GAG' for all full row rank $s \times n$ matrices G, where A is an $s \times s$ matrix and $s \leq n$.

Proof: $\forall y \in R^s$, $y \neq 0$, we have $y'GAG'y = z'Az > 0$, because G being of full row rank implies that $z = G'y \neq 0$ and A is positive definite. $\square$

Corollary 1: If A is symmetric and positive definite, so is A^{-1}.

Proof: Take G above to be A^{-1}. $\square$

Distribution of a Quadratic Form in Normal Random Variables

Suppose $Y \sim N(0, \sigma^2 I_n)$. The density of Y is $f_Y(y) = (2\pi\sigma^2)^{-n/2}\text{exp}(-y'y/(2\sigma^2))$. What is the distribution of $Y'AY$, where A is symmetric and positive semi-definite? (Sometimes we will be interested in $Y'AY$ when A is the inverse of a covariance matrix. In this case, A is symmetric and positive definite.) Answer: First write $B'AB = \Lambda$. Then, $Z = B'Y \sim N(0, \sigma^2 B'I_nB) = N(0, \sigma^2 I_n)$. So, $Y'AY = Y'BB'ABB'Y = Z'\Lambda Z = \sigma^2 \sum_{i=1}^n \lambda_i (Z_i/\sigma)^2$. We conclude that $Y'AY$ equals the sum of independent chi-squared random variables, each with one degree of freedom, scaled by certain constants. Now, the sum of independent chi-squared random variables is chi-squared with degrees of freedom given by the sum of the degrees of freedom. This is the reproductive property of chi-squared random variables. So, if λ_i equals 0 or 1 for all $i = 1, ..., n$, then $Y'AY/\sigma^2$ has a chi-squared distribution with degrees of freedom equal to $\sum_{i=1}^n \lambda_i$. This happens when A is a symmetric and idempotent matrix. So, we will now consider a special case of symmetric matrices called idempotent matrices.

Idempotent Matrices

A symmetric $n \times n$ matrix P for which $P^2 = P$ is called *idempotent*.

Lemma 8: The eigenvalues of a symmetric and idempotent matrix are zeros and/or ones and $rk(P) = \sum_{i=1}^n \lambda_i = \text{tr}(P)$, where $rk(P)$ denotes the rank of P.

Proof: $PB = B\Lambda$ (by the definition of eigenvalues and eigenvectors). Thus, we have $B\Lambda = PB = P^2B = PB\Lambda = B\Lambda\Lambda = B\Lambda^2$. So, $\Lambda = \Lambda^2$and $\lambda_j^2 = \lambda_j$, $\forall j = 1, ..., n$. That is, $\lambda_j = 0$ or 1, $\forall j$. The rank of a matrix equals its number of non-zero eigenvalues, so $rk(P) = \sum_{i=1}^n \lambda_i = \text{tr}(P)$. □

Lemma 9: If P is symmetric and idempotent, then P is positive semi-definite.

Proof: $Y'PY = (PY)'PY \geq 0$. □

Lemma 10: If P is symmetric and idempotent and $Y \sim N(0, \sigma^2 I_n)$, then $Y'PY/\sigma^2 \sim \chi^2_{rk[P]} = \chi^2_{\text{tr}[P]}$, where $\chi^2_{rk[P]}$ denotes a chi-squared distribution with $rk[P]$ degrees of freedom.

Proof: This follows from above. □

Comment: The converse of this theorem is also true. That is, if $Y \sim N(0, \sigma^2 I_n)$ and $Y'PY/\sigma^2 \sim \chi^2_{rk[P]}$, then P is symmetric and idempotent.

Lemma 11: If P is symmetric and idempotent and $Y \sim N(\mu, \sigma^2 I_n)$, then $Y'PY/\sigma^2 \sim \chi^2(rk[P], \delta)$, where the latter denotes a non-central chi-squared distribution with $rk[P]$ degrees of freedom and non-centrality parameter $\delta = \mu' P\mu/\sigma^2$.

Note that, by definition, if $Y \sim N(\mu, I_n)$, then $Y'Y \sim \chi^2(n, \mu'\mu)$. That is, a non-central chi-squared distribution is the distribution of a sum of squares of independent normal random variables with arbitrary means and unit variances. Why does this distribution depend only on $\mu'\mu$? Why does it not depend on each element $(\mu_1, ..., \mu_n)$ of μ? Think of a picture in R^n. The distribution of Y is spherically symmetric about μ.

Proof: Same as above. Write $B'PB = \Lambda$. Let $Z = B'Y$. Then, $Y'PY/\sigma^2 = Z'\Lambda Z/\sigma^2 = \sum_{i=1}^n \lambda_i Z_i^2/\sigma^2$, where $\lambda_i = 0$ or 1 and $Z \sim N(B'\mu, \sigma^2 B' I_n B) = N(B'\mu, \sigma^2 I_n)$. So, $Y'PY/\sigma^2$ has a non-central χ^2 distribution with degrees of freedom equal to $rk[P]$ and noncentrality parameter

$$\frac{(\Lambda B'\mu)'}{\sigma} \frac{(\Lambda B'\mu)}{\sigma} = \frac{\mu' B\Lambda^2 B'\mu}{\sigma^2} = \frac{\mu' B\Lambda B'\mu}{\sigma^2} = \frac{\mu' P\mu}{\sigma^2}. \qquad \square$$

Lemma 12: If P is idempotent, then so is $I_n - P$.

Proof: $(I_n - P)^2 = I_n^2 - I_n P - P I_n + P^2 = I_n - P$. □

Hence, $I_n - P$ is positive semi-definite, whenever P is idempotent.

Lemma 13: If P_1 and P_2 are idempotent, $P_1P_2 = 0$, and $Y \sim N(\mu, \sigma^2 I_n)$, then $Y'P_1Y$ and $Y'P_2Y$ are independent.

Proof: We can write $Y'P_1Y = (P_1Y)'P_1Y$ and $Y'P_2Y = (P_2Y)'(P_2Y)$. Because Y is multivariate normal, so are P_1Y and P_2Y. Their covariance matrix is $E(P_1U)(P_2U)' = EP_1UU'P_2 = \sigma^2 P_1P_2 = 0$, where $U = Y - \mu$. $\square$

Why does this result hold? The answer has to do with the geometric properties of idempotent matrices. To understand these properties we need to look at matrices as linear transformations from one Euclidean space to another.