MATH 531 Homework 6
Quadratic forms and the multivariate normal
February 24, 2025

Some background:

Let $\mathbf{z}$ be distributed $MN(0, I_m)$ then we know that $\mathbf{z}^T\mathbf{z} = \sum_{i=1}^m \mathbf{z}_i^2$ is distributed $\chi^2(m)$. – a sum of iid $N(0,1)$ RVs. (We also know that $\chi^2(m)$ is a member of the gamma distribution family with shape $2m$ and scale 2.)

Problem 1

Let M be an $n \times n$ projection matrix and let k be the dimension of the subspace that M projects onto.

1(a) For the eigendecomposition $M = UDU^T$ show that the diagonal elements of D must be either 0 or 1.

Proof. Since M is a projection matrix, then it is idempotent, that is $M^2 = M$. We can use the definition of M as its eigendecomposition to analyze further.

$$\begin{aligned} M^2 &= M \\ (UDU^T)(UDU^T) &= UDU^T \end{aligned}$$

The matrix U is orthogonal, so we can simplify this to

$$\begin{aligned} UDIDU^T &= UDU^T \\ UD^2U^T &= UDU^T \\ \implies D^2 &= D \end{aligned}$$

Since D is simply a diagonal matrix, we can rewrite this expression as

$$\mathrm{diag}(\sigma_1^2, ..., \sigma_n^2) = \mathrm{diag}(\sigma_1, ..., \sigma_n)$$

This directly implies that $\sigma_i^2 = \sigma_i$ for $i = 1, ..., n$. The only real solution that satisfies this criterion is if $\sigma_i = 1, 0$ for $i = 1, ..., n$. $\square$

This is provides an alternative proof that $tr(M) = k$.

1(b) Let U be an $n \times n$ orthonormal matrix and $\mathbf{z}$ be distributed $MN(0, I_n)$. Show that $U^T\mathbf{z}$ is also distributed $MN(0, I_n)$.

Proof. Assume $U^T z \sim MN(0, I_n)$. Then, $(U^T z)^T(U^T z)$ is distributed $\chi^2(n)$. We can expand the product to observe.

$$(U^T z)^T(U^T z) = z^T U U^T z = z^T I z = z^T z$$

We know from above that if $z \sim MN(0, I_n)$, then $z^T z \sim \chi^2(n)$. It is given that $z \sim MN(0, I_n)$, which implies that $U^T z \sim MN(0, I_n)$. $\square$

1(c) Let $\mathbf{z}$ be distributed $MN(0, I_n)$ show that $\mathbf{z}^T M\mathbf{z}$ is distributed $\chi^2(k)$.

Proof. Let $y = U^T z$. We know from the previous problem that $y \sim MN(0, I_n)$. Given that $M = UDU^T$, we can write $z^T M z$ as

$$\begin{aligned} z^T M z &= z^T(UDU^T)z \\ &= (U^T z)^T D U^T z \\ &= y^T D y = \sum_{i=1}^{n} \sigma_i y_i^2 \end{aligned}$$

We already know that $tr(M) = k$, meaning that there are exactly k elements of D that are equal to 1. Hence, we have

$$\sum_{i=1}^{n} \sigma_i y_i^2 = \sum_{i=1}^{k} y_i^2$$

Therefore, we can say that each y_i for $i = 1, ..., k$ is an i.i.d normally distributed random variable. Hence, $z^T M z = y^T D y = \sum_{i=1}^{k} y_i^2 \sim \chi^2(k)$. $\square$

Problem 2

Let $\mathbf{W}_1 = \mathbf{z}^T M\mathbf{z}$ and let $\mathbf{W}_2 = \mathbf{z}^T(I_n - M)\mathbf{z}$ with $\mathbf{z}$ be distributed $MN(0, I_n)$. Show that $\mathbf{W}_1$ and $\mathbf{W}_2$ are independent χ^2 RVs and explain why

$$F = \frac{\mathbf{W}_1/k}{\mathbf{W}_2/(n-k)}$$

has an F distribution with degrees of freedom (k, (n-k)).

This is a special case of the more general result:
If $\mathbf{z} \sim MN(0, I_n)$ and $AB = 0$ then $\mathbf{z}^T A\mathbf{z}$ are $\mathbf{z}^T B\mathbf{z}$ are independent RVs. You can use this for 4(b).

Proof. From (1c), we know that $z^T Mz \sim \chi^2(k)$ since M is a rank-k projection matrix (since it has k eigenvalues that are 1). Thus, the matrix $I_n - M$ has $n-k$ nonzero eigenvalues and is therefore rank$(n-k)$ and projects onto the complementary subspace that M projects onto. Now, let $A = M$ and $B = (I_n - M)$. Then, we have $W_1 = z^T Az$ $W_2 = z^T Bz$. By definition, if A and B are orthogonal matrices, then they satisfy $AB = 0$. We compute

$$AB = M(I_n - M) = M - M^2 = 0$$

since M is symmetric and idempotent. Clearly, A and B are orthogonal projection matrices, so it would follow that the matrices project onto orthogonal subspaces. Since z is a multivariate normal vector, then the products $z^T Az$ and $z^T Bz$ lie in orthogonal subspaces, and thus, we conclude that W_1 and W_2 are independent random variables.

Since $W_1 \sim \chi^2(k)$, then $W_1/k \sim \chi^2(k)/k$. The similar argument can be made for $W_2 \sim \chi^2(n-k)/(n-k)$. Therefore,

$$F = \frac{\chi^2(k)/k}{\chi^2(n-k)/(n-k)}$$

is an F-distribution with $(k, n-k)$ degrees of freedom. □

Problem 3

Consider the linear model

$$\mathbf{y} = X\beta + \mathbf{e}$$

with $\mathbf{e} \sim MN(0, \sigma^2 I_n)$, X with full rank and M the projection matrix onto $\mathcal{W}_{\mathcal{X}}$. X has k columns.

The twist:
Partition the regression matrix as

$$X = [X_1|X_2]$$

with X_1 having j columns (X_2 having $k - j$) and M_1 the projection matrix onto $\mathcal{W}_{X_1}$. Also partition $\beta = [\beta_1, \beta_2]$.

3(a) Explain why $M - M_1$ is also a projection matrix and identify its subspace.

We know that M projects onto $\mathcal{W}_X$, and M_1 projects onto $\mathcal{W}_{X_1}$. By subtracting $M - M_1$, then this matrix essentially projects onto a subspace of $\mathcal{W}_X$ such that it is not in the subspace $\mathcal{W}_{X_1}$. Simply put, $(M - M_1)$ projects onto the subspace of W_X that is spanned by the columns of X_2.

3(b) Explain why $(1/\sigma^2)\mathbf{y}^T(I - M)\mathbf{y}$ is distributed $\chi^2(n - k)$.

Proof. First, we can left-multiply the linear model expression by the projection matrix $(I - M)$.

$$(I - M)y = (I - M)(X\beta + e)$$

Since $X\beta \in \mathcal{W}_X$, then that leaves us with

$$(I - M)y = (I - M)e$$

In quadratic form, we have

$$y^T(I - M)y = e^T(I - M)e$$

We can normalize e by dividing out the expression by σ^2. We have

$$\frac{1}{\sigma^2}y^T(I-M)y = \frac{1}{\sigma^2}e^T(I-M)e$$

Let $\varepsilon = \frac{1}{\sigma^2}e$. Since $e \sim MN(0, \sigma^2 I_n)$, then the normalized vector $\varepsilon \sim MN(0, I_n)$. From (2), we know that for a random vector $z \sim MN(0, I_n)$, the for a rank-k projection matrix M, then $z^T(I-M)z \sim \chi^2(n-k)$. Therefore, we can say that $\frac{1}{\sigma^2}e^T(I-M)e \sim \chi^2(n-k)$, and equivalently, $\frac{1}{\sigma^2}y^T(I-M)y \sim \chi^2(n-k)$. $\square$

3(c) *The classic ANOVA decomposition*
Show that

$$\mathbf{y}^T\mathbf{y} = \mathbf{y}^T M_1\mathbf{y} + \mathbf{y}^T(M-M_1)\mathbf{y} + \mathbf{y}^T(I-M)\mathbf{y}$$

or equivalently

$$\mathbf{y}^T(I-M_1)\mathbf{y} = \mathbf{y}^T(M-M_1)\mathbf{y} + \mathbf{y}^T(I-M)\mathbf{y}$$

Proof. Through some simple algebra, we can say that $I = M_1 + (M - M_1) + (I - M)$. Therefore, we can say that

$$y^Ty = y^T(M_1 + (M-M_1) + (I-M))y$$

We can algebraically simplify this expression to be

$$y^Ty = y^T M_1 y + y^T(M-M_1)y + y^T(I-M)y$$

Furthermore, we can subtract the LHS by the first term on the RHS to find that

$$\begin{aligned} y^Ty - y^T M_1 y &= y^T(M-M_1)y + y^T(I-M)y \\ \to y^T(I-M_1)y &= y^T(M-M_1)y + y^T(I-M)y \quad \square \end{aligned}$$

Problem 4

4(a) Show that $(1/\sigma^2)\mathbf{y}^T(M-M_1)\mathbf{y}$ is $\chi^2(k-j)$ when $\beta_2 = 0$.

Hint: We know this is true for $(1/\sigma^2)\mathbf{e}^T(M - M_1)\mathbf{e}$ the main task is to show that the mean of $\mathbf{y}$ is canceled by $M - M_1$.

Proof. Given that $\beta_2 = 0$, the model simply becomes

$$y = X_1\beta_1 + e$$

So the "mean part" of y $(X_1\beta_1)$ is projected onto $\mathcal{W}_{X_1}$, which is orthogonal to the column space spanned by $(M - M_1)$, that is $(M - M_1)(X_1\beta_1) = 0$. We use this fact to observe:

$$y^T(M - M_1)y = (X_1\beta_1 + e)^T(M - M_1)(X_1\beta_1 + e)$$

Thus, the mean part contributes nothing to the model and this simply becomes a white noise process. We have

$$y^T(M - M_1)y = e^T(M - M_1)e$$

We can again normalize this expression by dividing it by the variance.

$$\frac{1}{\sigma^2}y^T(M - M_1)y = \frac{1}{\sigma^2}e(M - M_1)e$$

Again, let $\varepsilon = \frac{1}{\sigma^2}e$. Since $(M - M_1)$ has rank$(k - j)$, then we have that $\frac{1}{\sigma^2}e^T(M - M_1)e \sim \chi^2(k - j)$, and therefore

$$\frac{1}{\sigma^2}y^T(M - M_1)y \sim \chi^2(k - j)$$

4(b) Show that $\mathbf{y}^T(M - M_1)\mathbf{y}$ and $\mathbf{y}^T(I - M)\mathbf{y}$ are independent.

Proof. To show that these terms are independent, we must show that their constituent projection matrices are orthogonal, that is, $(M - M_1)(I - M) = 0$. We start by expanding this expression.

$$(M - M_1)(I - M) = M - M^2 - M_1 + M_1 M$$

M is idempotent, so we can say that

$$M - M^2 - M_1 + M_1 M = -M_1 + M_1 M$$

Now, since M and M_1 are both square, symmetric, and idempotent, and $\mathcal{W}_{X_1} \subseteq \mathcal{W}_X$, then we can say that $M_1 M = M_1$, since projecting M onto $\mathcal{W}_{X_1}$ involves only the projection matrix M_1. Therefore, we have

$$(M - M_1)(I - M) = 0$$

Since the respective projection matrices are orthogonal, then we can say that they project onto orthogonal subspaces. Therefore, we can say that $y^T(M - M_1)y$ and $y^T(I - M)y$ are independent of each other. $\square$

This is the basic ingredient to justify the usual F test for testing whether a subset of parameters (β_2 in this case) is zero.