MATH 531 Homework 6
Quadratic forms and the multivariate normal
February 20, 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.

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 $MN(0, I_n)$.

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

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).

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.

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

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

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

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 $M - M_1$.

4(b) Show that $\mathbf{y}^T(M - M_1)\mathbf{y}$ and $\mathbf{y}^T(I - M)\mathbf{y}$ are independent. Does this depend on the condition $\beta_2 = 0$?

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