Econometrics 3 (ECOM90013) Assignment 1

1. **In this question we will assume that $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a p-vector, as is $\boldsymbol{\mu}$, and that the $\boldsymbol{p}$ x $\boldsymbol{p}$ matrix $\boldsymbol{\Sigma} > \mathbf{0}$.**

    a. **(1 mark) If v is any fixed p-vector, show that**

$$\boldsymbol{g} = \frac{\boldsymbol{v}'(\mathbf{x} - \boldsymbol{\mu})}{\sqrt{\boldsymbol{v}'\boldsymbol{\Sigma}\boldsymbol{v}}} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{1})$$

The variable g represents a standard normal version of x which has had a linear transformation applied to it. To achieve this form, subtract the mean from x, apply the linear transformation and then divide by its standard deviation.

More formally, given:

$$x \sim N(\mu, \Sigma)$$

$$x - \mu \sim N(0, \Sigma)$$

Then, after applying the fixed p-vector v:

$$v'(x - \mu) \sim N(0, v'\Sigma v)$$

This variance is given by the definition of the variance of a linear transformation ($Var(Ax) = A'Var(x)A$). We then normalise by the standard deviation to derive g:

$$g = \frac{v'(\mathrm{x} - \mu)}{\sqrt{v'\Sigma v}} \sim N(0,1)$$

To verify the expected mean:

$$E[g] = E[v'(x - \mu)] = v'(E[x] - \mu) = v'(\mu - \mu) = 0$$

To verify the variance (using the same variance of a linear transformation definition as above for the numerator):

$$Var(g) = \frac{Var(v'(\mathrm{x} - \mu))}{Var(\sqrt{v'\Sigma v})} = \frac{v'(\mathrm{Var}(\mathrm{x} - \mu))}{v'\Sigma v} = \frac{v'\Sigma v}{v'\Sigma v} = 1$$

Therefore:

$$g = \frac{v'(\mathrm{x} - \mu)}{\sqrt{v'\Sigma v}} \sim N(0,1)$$

Josh Copeland (SID: 1444772)

b. (4 marks) If v is now a random vector independent *of x* for which $\mathrm{P}(\boldsymbol{v}'\boldsymbol{\Sigma v} = \mathbf{0}) = \mathbf{0}$, show that $\boldsymbol{g} \sim N(\mathbf{0}, \mathbf{1})$ *and is independent of v. Why have we assumed* $\mathrm{P}(\boldsymbol{v}'\boldsymbol{\Sigma v} = \mathbf{0}) = \mathbf{0}$? Provide an equivalent statement of this assumption not involving $\boldsymbol{\Sigma}$.

Although v is now random, it is independent of x by definition, meaning that this numerator can be expressed as a linear combination of independent normal random variables. This means we can condition x on v, as the latter's randomness adds nothing to our understanding of the prior – which means it works as if a linear transformation.

Therefore, due to the assumption of independence between v and x, we can prove $g{\sim}N(0{,}1)$ using the same proofs as above in Question 1a. given that:

$$g = g|v = \frac{v'(\mathrm{x} - \mathrm{\mu})}{\sqrt{v'\Sigma v}} \sim N(0{,}1)$$

We can prove independence by leveraging this knowledge about the conditional distribution of g.

Given v, we know that:

$$v'(\mathrm{x} - \mathrm{\mu}) = v'(x - \mu)|v \sim N(0, v'\Sigma v)$$

$v'(\mathrm{x} - \mathrm{\mu})$ is simply just a linear combination of normally distributed random variables. Therefore, g's numerator must be independent of v.

Similarly, g's denominator is given by $\sqrt{v'\Sigma v}$, which depends only on v, and hence is not random one v is known. It by definition does not depend on g.

We have assumed $\mathrm{P}(v'\Sigma v = 0) = 0$ because this guarantees that the denominator of g is non-zero, ensuring this variable is always well-defined. This could alternatively be expressed as:

$$P(v'v = 0) = 0$$

Implying that the quadratic form of v is zero only when v is the zero vector.

***c. (2 marks) Hence, show that if $\boldsymbol{y} = [\boldsymbol{y_1}, \boldsymbol{y_2}, \boldsymbol{y_3}]' \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{I_3})$*, then**

$$\boldsymbol{h} = \frac{\boldsymbol{y_1 e^{y_3} + y_2 log|y_3|}}{\boldsymbol{[e^{2y_3} + (log|y_3|)^2]^{1/2}}} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{1})$$

First, find the expected mean:

$$E[h] = \frac{E[y_1 e^{y_3} + y_2 \log|y_3|]}{E[[e^{2y_3} + (\log|y_3|)^2]^{1/2}]}$$

$$E[h] = \frac{E[y_1] e^{y_3} + E[y_2] \log|y_3|]}{E[[e^{2y_3} + (\log|y_3|)^2]^{1/2}]}$$

Given $y = [y_1, y_2, y_3]' \sim N(0, I_3)$:

$$E[h] = \frac{0.\, e^{y_3} + 0.\log|y_3|]}{E[[e^{2y_3} + (\log|y_3|)^2]^{1/2}]} = 0$$

Now we begin finding the variance by first evaluating h^2:

$$h^2 = \frac{(e^{y_3} + y_2 \log|y_3|)^2}{e^{2y_3} + (\log|y_3|)^2}$$

$$h^2 = \frac{e^{2y_3} + y_2^2 (\log|y_3|)^2 + 2y_1 y_2 e^{y_3} \log|y_3|}{e^{2y_3} + (\log|y_3|)^2}$$

Because we know the elements of y are independent of each other, their covariances cancel out as they are equal to zero by definition, meaning we only need to keep the squared terms in our results when we take expectations:

$$E[h^2] = \frac{E[e^{2y_3} + y_2^2 (\log|y_3|)^2 + 2y_1 y_2 e^{y_3} \log|y_3|]}{E[e^{2y_3} + (\log|y_3|)^2]}$$

$$E[h^2] = \frac{E[e^{2y_3} + 1(\log|y_3|)^2 + 0]}{E[e^{2y_3} + (\log|y_3|)^2]} = \frac{e^{2y_3} + (\log|y_3|)^2}{e^{2y_3} + (\log|y_3|)^2} = 1$$

Therefore,

$$h = \frac{y_1 e^{y_3} + y_2 \log|y_3|}{[e^{2y_3} + (log|y_3|)^2]^{\frac{1}{2}}} \sim N(0,1)$$

2. **(3 marks) Suppose that $\boldsymbol{x} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where the $\boldsymbol{p}$ x $\boldsymbol{p}$ matrix $\boldsymbol{\Sigma} > \mathbf{0}$, and the $\boldsymbol{v}$ is a fixed $\boldsymbol{p}$-vector. If $\boldsymbol{r_i}$, the $\boldsymbol{i}$th element of the vector $\boldsymbol{r}$, is the correlation between $\boldsymbol{x_i}$ and $\boldsymbol{v'x}$, show that $\boldsymbol{r} = (\boldsymbol{cD})^{-1/2} \boldsymbol{\Sigma v}$, where $\boldsymbol{c} = \boldsymbol{v' \Sigma v}$ and $\boldsymbol{D} = \boldsymbol{diag}(\boldsymbol{\Sigma})$.**

We are told that the i-th element of the vector r's correlation:

$$r_i = \frac{Cov(x_i, v'x)}{\sqrt{Var(x_i)}\sqrt{Var(v'x_i)}} = (cD)^{-1/2}\Sigma v$$

Therefore, this result should be self-evident after deriving each component.

We can immediately simplify this by substituting in $c = v'\Sigma v$, which is the variance of the linear transformation of x, $v'\Sigma v$ (as established in Question 1a). Therefore:

$$r_i = \frac{Cov(x_i, v'x)}{\sqrt{Var(x_i)}\sqrt{c}}$$

The variance of x_i is also given in the question. Given:

$$x \sim N(\mu, \Sigma)$$

It must be that:

$$Var(x_i) = \Sigma_{ii} = D$$

Therefore:

$$r_i = \frac{Cov(x_i, v'x)}{\sqrt{D}\sqrt{c}} = (cD)^{-\frac{1}{2}}Cov(x_i, v'x)$$

Now to find the covariance (noting that $Cov(A, B) = E[AB] - E[A]E[B]$):

$$Cov(x_i, v'x) = E[x(v'x)'] - E[x]E[v'x]$$

Simplifying this further:

$$E[x(v'x)'] - E[x]E[v'x] = E[xx']v - E[x]v'E[x]$$

Collecting terms and noting the definition of Σ:

$$E[xx']v - E[x]v'E[x] = (\Sigma + \mathrm{E}[x]E[x'])v - E[x]v'E[x]$$

$$= \Sigma v + \mathrm{E}[x]E[x']v - \mathrm{E}[x]v'E[x]$$

Because v is scalar, we can write:

$$\mathrm{E}[x]v'E[x'] = (\mathrm{E}[x]v'E[x'])v$$

Therefore, when we apply this and take the difference between the second and third terms we have:

$$Cov(x_i, v'x) = \Sigma v + (\mathrm{E}[x]E[x'])v - (\mathrm{E}[x]v'E[x'])v = \Sigma v$$

Therefore, by plugging this back into to what we're given, we can conclude:

$$r_i = \frac{Cov(x_i, v'x)}{\sqrt{Var(x_i)}\sqrt{Var(v'x_i)}} = (cD)^{-1/2}\Sigma v$$

(1 mark) Bonus question: when does $r = \Sigma v$?

$r = \Sigma v$ holds when $(cD)^{-1/2}$ is the identity matrix. This would imply that the scaling factor $(cD)^{-1/2}$ is the identity, meaning no scaling is applied to Σv.