

# Xu\_Hong\_solutions\_HW3

Hong Xu

September 17, 2015

## 1.

Prove that the total variance of  $\mathbf{X}$  is the MSE of the mean:

$$E\|\mathbf{X} - E(\mathbf{X})\|^2 = \text{tr}(\text{Var}(\mathbf{X})).$$

Proof:

We can show that the MSE of the mean:  $E\|\mathbf{X} - E(\mathbf{X})\|^2 = E(\sum_i (X_i - E(X_i))^2) = \sum_i E((X_i - E(X_i))^2)$

The covariance of  $\mathbf{X}$ :  $\text{Var}(\mathbf{X}) = E((\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^T)$

So the total variance of  $\mathbf{X}$  equals to  $\text{tr}(\text{Var}(\mathbf{X})) = \sum_i E((X_i - E(X_i))(X_i - E(X_i))^T) = \sum_i E((X_i - E(X_i))^2)$ , which equals to the MSE of the mean.

## 2.

Suppose  $\mathbf{Y} \sim N_n(0, I)$  and let  $A$  be symmetric. Show that

$$\mathbf{Y}^T A \mathbf{Y} \stackrel{\text{dist}}{=} \sum_i \lambda_i W_i,$$

where  $W_i \stackrel{iid}{\sim} \chi_1^2$  and  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of  $A$ .

Proof:

Since  $A$  is symmetric, we have the spectral decomposition of  $A = P D P^T$ ,

where  $P$  is orthogonal,  $D$  is a diagonal matrix  $\text{diag}(\lambda_1, \dots, \lambda_n)$

So  $\mathbf{Y}^T A \mathbf{Y} = \mathbf{Y}^T P D P^T \mathbf{Y} = (P^T \mathbf{Y})^T D (P^T \mathbf{Y})$

and  $P^T \mathbf{Y} \sim N_n(P^T 0, P^T I P)$ , since  $P$  is orthogonal, we have  $P^T \mathbf{Y} \sim N_n(0, I)$ .

With that, we can show  $\mathbf{Y}^T A \mathbf{Y} = \mathbf{Y}^T D \mathbf{Y} = \sum_i \sum_j \lambda_{ij} Y_i Y_j = \sum_i \lambda_i Y_i^2$ .

And  $\mathbf{Y} \sim N_n(0, I)$ ,

So,  $Y_i \sim N_n(0, 1)$ ,  $W_i = Y_i^2 \sim \chi_1^2$ .

So that  $\mathbf{Y}^T A \mathbf{Y} \stackrel{\text{dist}}{=} \sum_i \lambda_i W_i$ .

## 3.

a)

Show that this model implies that there is a positive covariance for within-state poverty rates over time,  $\text{Cov}[Y_{i1}, Y_{i2}] = \text{Var}[U_i] = \sigma_u^2$ .

Proof:

Since  $U_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_u^2)$  and  $\epsilon_{it} \stackrel{ind}{\sim} \mathcal{N}(0, \sigma_{\epsilon_t}^2)$ ,

and  $Y_{it} = \mu_t + U_i + \epsilon_{it}$ ,

Then, at time  $t$ ,  $E(Y_{it}) = \mu_t + E(U_i) + E(\epsilon_{it}) = \mu_t + 0 + 0 = \mu_t$ .

So,  $\text{Cov}[Y_{i1}, Y_{i2}] = E((Y_{i1} - E(Y_{i1}))(Y_{i2} - E(Y_{i2}))) = E((Y_{i1} - \mu_1)(Y_{i2} - \mu_2)) = E((U_i + \epsilon_{i1})(U_i + \epsilon_{i2})) = E(U_i^2 + \epsilon_{i1}U_i + \epsilon_{i2}U_i + \epsilon_{i1}\epsilon_{i2}) = E(U_i^2) + E(\epsilon_{i1}U_i) + E(\epsilon_{i2}U_i) + E(\epsilon_{i1}\epsilon_{i2})$ .

Since  $U_i$  and  $\epsilon_{it}$  are independent, it's easy to show the last three expectations are zero. So,  $\text{Cov}[Y_{i1}, Y_{i2}] = E(U_i^2)$

Also,  $\text{Var}(U_i) = E(U_i^2) - E(U_i)^2 = E(U_i^2)$ ,

So we have  $\text{Cov}[Y_{i1}, Y_{i2}] = E(U_i^2) = \text{Var}(U_i) = \sigma_\mu^2$

**b)**

Answer:

Since  $Y_{it} = \mu_t + U_i + \epsilon_{it}$ ,

$Y_{it} \sim N(\mu_t, \sigma_\mu^2 + \sigma_{\epsilon_t}^2)$ .

Obviously  $Y_{i1}$  and  $Y_{i2}$  follow normal distribution, so that  $\mathbf{Y}_i = (Y_{i1}, Y_{i2})^T$  would be an  $\mathcal{MVN}$  distribution.

Its mean vector  $\mu$  would be  $E(\mathbf{Y}_i) = (E(Y_{i1}), E(Y_{i2}))^T = (\mu_1, \mu_2)^T$ , given the proof we showed in (a).

Its  $\Sigma$  would be:

$$\Sigma = \begin{bmatrix} \text{Var}(Y_{i1}) & \text{Cov}(Y_{i1}, Y_{i2}) \\ \text{Cov}(Y_{i2}, Y_{i1}) & \text{Var}(Y_{i2}) \end{bmatrix}$$

Given the proof in a)  $\text{Cov}[Y_{i1}, Y_{i2}] = E(U_i^2) = \text{Var}(U_i) = \sigma_\mu^2$ ,

$$\Sigma = \begin{bmatrix} \sigma_\mu^2 + \sigma_{\epsilon_1}^2 & \sigma_\mu^2 \\ \sigma_\mu^2 & \sigma_\mu^2 + \sigma_{\epsilon_2}^2 \end{bmatrix}$$

**c)**

```
# skip the first row
saipe98 <- read.csv("CPS98T.TXT", skip = 1, sep="")
saipe05 <- read.csv("CPS05T.TXT", skip = 1, sep="")
# for children aged 5-17, CPS poverty is at column cps98.1 or cps05.1
cps98 <- saipe98[, "cps98.1"]
cps05 <- saipe05[, "cps05.1"]
# construct a data frame
cps <- cbind(cps98, cps05)
```

**d)**

```
## find the mean poverty rate
colMeans(cps)
```

```
##      cps98      cps05
## 16.97407 15.69192
```

e)

```
## find the covariance matrix for Y, i.e. cps
cov(cps)
```

```
##          cps98    cps05
## cps98 39.92843 27.59969
## cps05 27.59969 33.05264
```

f)

In b) we showed:

$$\Sigma = \begin{bmatrix} \sigma_{\mu}^2 + \sigma_{e_1}^2 & \sigma_{\mu}^2 \\ \sigma_{\mu}^2 & \sigma_{\mu}^2 + \sigma_{e_2}^2 \end{bmatrix}$$

and

$$\mu = (\mu_1, \mu_2)^T$$

So,

$$\mu_1 = 16.97407$$

$$\mu_2 = 15.69192$$

$$\sigma_{\mu}^2 = 27.59969$$

$$\sigma_{e_1}^2 = 39.92843 - \sigma_{\mu}^2 = 39.92843 - 27.59969 = 12.32874$$

$$\sigma_{e_2}^2 = 33.05264 - \sigma_{\mu}^2 = 33.05264 - 27.59969 = 5.45295$$

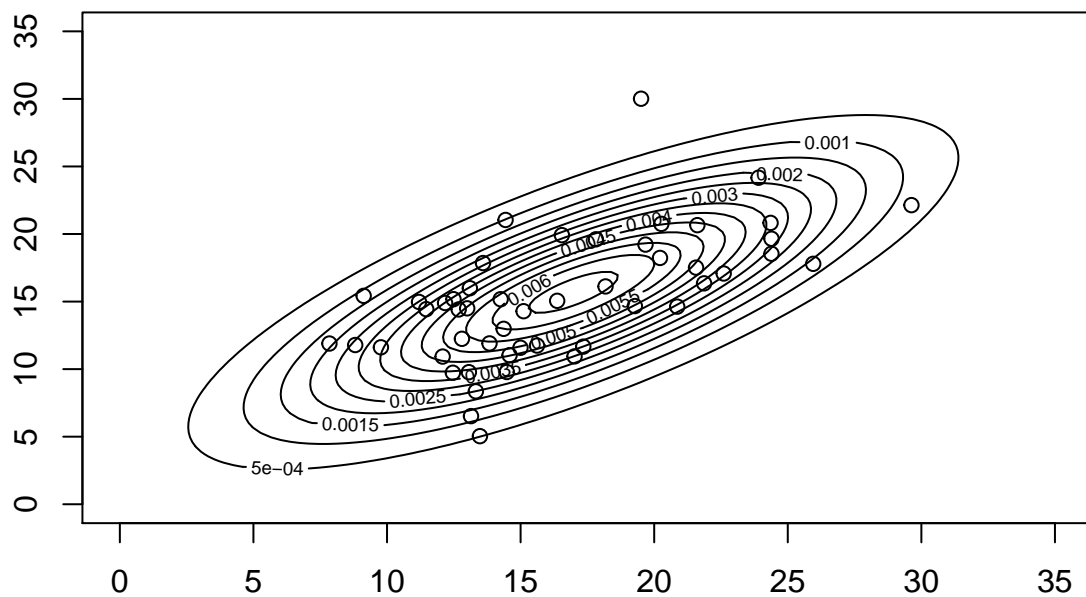
g)

```
require(mvtnorm)
```

```
## Loading required package: mvtnorm
```

```
x.points <- seq(0, 35, length.out = 100)
y.points <- x.points
z <- matrix(0, 100, 100)
mu <- colMeans(cps)
sigma <- cov(cps)

for (i in 1:100) {
  for (j in 1:100) {
    z[i, j] <- dmvnorm(c(x.points[i], y.points[j]), mean = mu, sigma = sigma)
  }
}
contour(x.points, y.points, z)
# add points
points(cps[,1], cps[,2])
```



```
## The multivariate normal looks like a good fit to the data
## given the density and the rotation the data points appear.
```