Lecture 15

15.1 Orthogonal transformation of standard normal sample.

Consider $X_1, \ldots, X_n \sim N(0,1)$ i.d.d. standard normal r.v. and let V be an orthogonal transformation in $\mathbb{R}^n$. Consider a vector $\vec{Y} = \vec{X}V = (Y_1, \ldots, Y_n)$. What is the joint distribution of $Y_1, \ldots, Y_n$? It is very easy to see that each Y_i has standard normal distribution and that they are uncorrelated. Let us check this. First of all, each

$$Y_i = \sum_{k=1}^{n} v_{ki} X_k$$

is a sum of independent normal r.v. and, therefore, Y_i has normal distribution with mean 0 and variance

$$\mathrm{Var}(Y_i) = \sum_{k=1}^{n} v_{ik}^2 = 1,$$

since the matrix V is orthogonal and the length of each column vector is 1. So, each r.v. $Y_i \sim N(0,1)$. Any two r.v. Y_i and Y_j in this sequence are uncorrelated since

$$\mathbb{E} Y_i Y_j = \sum_{k=1}^{n} v_{ik} v_{jk} = \vec{v}_i' \vec{v}_j' = 0$$

since the columns $\vec{v}_i' \perp \vec{v}_j'$ are orthogonal.

Does uncorrelated mean independent? In general no, but for normal it is true which means that we want to show that Y's are i.i.d. standard normal, i.e. $\vec{Y}$ has the same distribution as $\vec{X}$. Let us show this more accurately. Given a vector $t = (t_1, \ldots, t_n)$, the moment generating function of i.i.d. sequence $X_1, \ldots, X_n$ can be computed as follows:

$$\varphi(t) = \mathbb{E} e^{\vec{X} t^T} = \mathbb{E} e^{t_1 X_1 + \ldots + t_n X_n} = \prod_{i=1}^{n} \mathbb{E} e^{t_i X_i}$$

$$= \prod_{i=1}^{n} e^{\frac{t_i^2}{2}} = e^{\frac{1}{2}\sum_{i=1}^{n} t_i^2} = e^{\frac{1}{2}|t|^2}.$$

On the other hand, since $\vec{Y} = \vec{X}V$ and

$$t_1 Y_1 + \ldots + t_n Y_n = (Y_1, \ldots, Y_n) \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} = (Y_1, \ldots, Y_n) t^T = \vec{X} V t^T,$$

the moment generating function of $Y_1, \ldots, Y_n$ is:

$$\mathbb{E} e^{t_1 Y_1 + \cdots + t_n Y_n} = \mathbb{E} e^{\vec{X} V t^T} = \mathbb{E} e^{\vec{X}(tV^T)^T}.$$

But this is the moment generating function of vector $\vec{X}$ at the point tV^T, i.e. it is equal to

$$\varphi(tV^T) = e^{\frac{1}{2}|tV^T|^2} = e^{\frac{1}{2}|t|^2},$$

since the orthogonal transformation preserves the length of a vector $|tV^T| = |t|$. This means that the moment generating function of $\vec{Y}$ is exactly the same as of $\vec{X}$ which means that $Y_1, \ldots, Y_n$ have the same joint distribution as X's, i.e. i.i.d. standard normal.

Now we are ready to move to the main question we asked in the beginning of the previous lecture: What is the joint distribution of $\bar{X}$ (sample mean) and $\bar{X^2} - (\bar{X})^2$ (sample variance)?

Theorem. *If $X_1, \ldots, X_n$ are i.d.d. standard normal, then sample mean $\bar{X}$ and sample variance $\bar{X^2} - (\bar{X})^2$ are independent,*

$$\sqrt{n}\bar{X} \sim N(0,1) \text{ and } n(\bar{X^2} - (\bar{X})^2) \sim \chi^2_{n-1},$$

i.e. $\sqrt{n}\bar{X}$ has standard normal distribution and $n(\bar{X^2} - (\bar{X})^2)$ has χ^2_{n-1} distribution with $(n-1)$ degrees of freedom.

Proof. Consider a vector $\vec{Y}$ given by transformation

$$\vec{Y} = (Y_1, \ldots, Y_n) = \vec{X} V = (X_1, \ldots, X_n) \begin{pmatrix} \frac{1}{\sqrt{n}} & \cdots & \cdots & \cdots \\ \vdots & \cdots & ? & \cdots \\ \frac{1}{\sqrt{n}} & \cdots & \cdots & \cdots \end{pmatrix}.$$

Here we chose a first column of the matrix V to be equal to

$$\vec{v}_1 = \Big(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\Big).$$

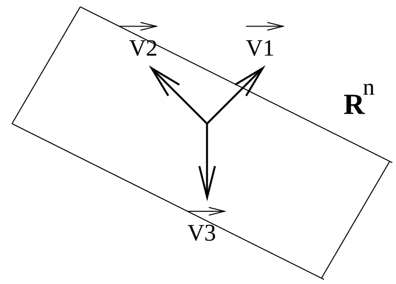

Figure 15.1: Unit Vectors.

We let the remaining columns be any vectors such that the matrix V defines orthogonal transformation. This can be done since the length of the first column vector $|\vec{v}_1| = 1$, and we can simply choose the columns $\vec{v}_2, \ldots, \vec{v}_n$ to be any orthogonal basis in the hyperplane orthogonal to vector $\vec{v}_1$, as shown in figure 15.1.

Let us discuss some properties of this particular transformation. First of all, we showed above that $Y_1, \ldots, Y_n$ are also i.i.d. standard normal. Because of the particular choice of the first column $\vec{v}_1$ in V, the first r.v.

$$Y_1 = \frac{1}{\sqrt{n}} X_1 + \ldots + \frac{1}{\sqrt{n}} X_n,$$

and, therefore,

$$\bar{X} = \frac{1}{\sqrt{n}} Y_1. \tag{15.1}$$

Next, n times sample variance can be written as

$$\begin{aligned} n(\bar{X^2} - (\bar{X})^2) &= X_1^2 + \ldots + X_n^2 - \Big(\frac{1}{\sqrt{n}}(X_1 + \ldots + X_n)\Big)^2 \\ &= X_1^2 + \ldots + X_n^2 - Y_1^2. \end{aligned}$$

But the orthogonal transformation V preserves the length

$$Y_1^2 + \cdots + Y_n^2 = X_1^2 + \cdots + X_n^2$$

and, therefore, we get

$$n(\bar{X^2} - (\bar{X})^2) = Y_1^2 + \ldots + Y_n^2 - Y_1^2 = Y_2^2 + \ldots + Y_n^2. \tag{15.2}$$

Equations (15.1) and (15.2) show that sample mean and sample variance are independent since Y_1 and $(Y_2, \ldots, Y_n)$ are independent, $\sqrt{n}\bar{X} = Y_1$ has standard normal distribution and $n(\bar{X^2} - (\bar{X})^2)$ has χ^2_{n-1} distribution since $Y_2, \ldots, Y_n$ are independent standard normal.

□

Consider now the case when

$$X_1, \ldots, X_n \sim N(\alpha, \sigma^2)$$

are i.i.d. normal random variables with mean α and variance σ^2. In this case, we know that

$$Z_1 = \frac{X_1 - \alpha}{\sigma}, \cdots, Z_n = \frac{X_n - \alpha}{\sigma} \sim N(0, 1)$$

are independent standard normal. Theorem applied to $Z_1, \ldots, Z_n$ gives that

$$\sqrt{n}\bar{Z} = \sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\frac{X_i - \alpha}{\sigma} = \frac{\sqrt{n}(\bar{X} - \alpha)}{\sigma} \sim N(0, 1)$$

and

$$\begin{aligned} n(\bar{Z^2} - (\bar{Z})^2) &= n\Bigl(\frac{1}{n}\sum\Bigl(\frac{X_i - \alpha}{\sigma}\Bigr)^2 - \Bigl(\frac{1}{n}\sum\frac{X_i - \alpha}{\sigma}\Bigr)^2\Bigr) \\ &= n\frac{1}{n}\sum_{i=1}^{n}\Bigl(\frac{X_i - \alpha}{\sigma} - \frac{1}{n}\sum\frac{X_i - \alpha}{\sigma}\Bigr)^2 \\ &= n\frac{\bar{X^2} - (\bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}. \end{aligned}$$