

Lectures of June 20th, 2006

Scribe: Min Shi, Vivekananthan Suganthan

1 Jointly Gaussian Random Variables

We will use \vec{x} denote $(x_1, x_2, \dots, x_n)^T$. Random variables $\vec{X} = (X_1, X_2, \dots, X_n)^T$ is said to jointly Gaussian when the jointly PDF $f_{\vec{x}}(\vec{x})$ takes the following form:

$$f_{\vec{x}}(\vec{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} |k|^{\frac{1}{2}}} e^{-\frac{1}{2}(\vec{x}-\vec{m})k^{-1}(\vec{x}-\vec{m})}$$

Where k is a $N \times N$ matrix (called the covariance matrix); and $\vec{m} = (m_1, m_2, \dots, m_n)^T$ with $m_i = E[x_i]$.

In particular:

$$k(i, j) = \text{cov}[x_i, x_j]$$

$$\text{note: } \text{cov}(x_i, x_i) = \text{var}[x_i].$$

It can be verified that when $n = 1$, the PDF is the Gaussian PDF is the Gaussian PDF for single Gaussian random variable.

Now consider the case $\vec{x} = (x_1, x_2)$

Case 1: $\sigma_1 = \sigma_2$

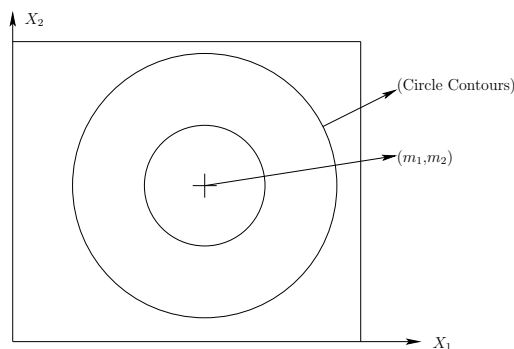


Figure 1: $\sigma_1 = \sigma_2$.

$$K = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \quad \text{i.e.} \quad \text{cov}(x_1, x_2) = 0$$

Case 2: $\sigma_1 \neq \sigma_2$

Condition 1: $\sigma_1 > \sigma_2$.

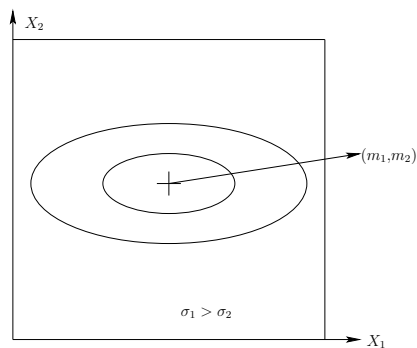


Figure 2: $\sigma_1 > \sigma_2$.

Condition 2: $\sigma_1 < \sigma_2$.

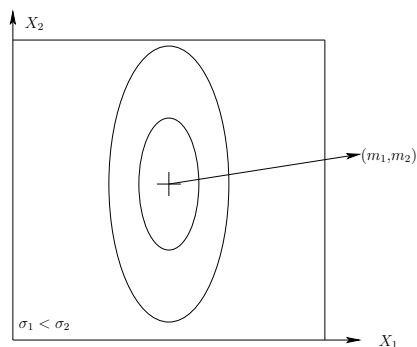


Figure 3: $\sigma_1 < \sigma_2$.

For case 2 with both conditions:

$$K = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \quad \sigma_1 \neq \sigma_2$$

Case 3: x_1 and x_2 are not independent

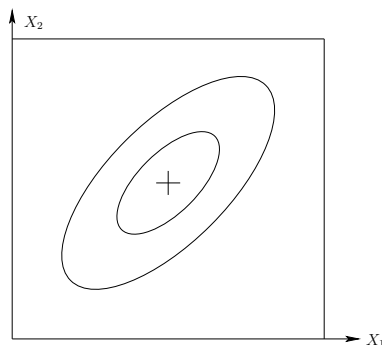


Figure 4: x_1 and x_2 not independent

$$K = \begin{bmatrix} \text{var}[x_1] & \text{cov}[x_1, x_2] \\ \text{cov}[x_2, x_1] & \text{var}[x_2] \end{bmatrix} \quad \text{No entry is zero.}$$

2 General Remarks on Jointly Gaussian Random Variables

1. If k is diagonal matrix, then X_1 and X_2 are independent (case 1 and case 2). That is, if two random variables are jointly Gaussian, then uncorelatedness and independence are equivalent.
2. If several random variable are jointly Gaussian, the each of them is Gaussian.

But, if two random variable are both Gaussian, they may not be jointly Gaussian.

Below is a counter example:

Let (x_1, x_2) be jointly Gaussian with $\vec{m} = (0, 0)$,

$$K = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Let $f_{x_1, x_2}(x_1, x_2)$ be the jointly PDF of (x_1, x_2) . We Define:

$$f_{y_1, y_2}(x_1, x_2) = \begin{cases} 2 \cdot f_{x_1, x_2}(x_1, x_2), & \text{if } x_1 \cdot x_2 > 0, \\ 0, & \text{otherwise.} \end{cases}$$

We can verify $f_{y_1, y_2}(x_1, x_2)$ is a PDF (i.e. integrated to 1). The random variable (y_1, y_2) following distribution f_{y_1, y_2} are not join Gaussian, but y_1, y_2 are Gaussian.

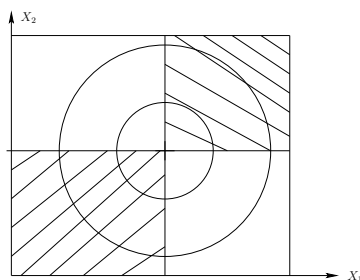


Figure 5: Figure of the Counter Example.

3 Sample mean, laws of large number and central limit theorem

Let x_1, x_2, \dots, x_n be n independent draw of random variable x . (That is x_1, x_2, \dots, x_n are independently identically distributed)(iid).

Denote: $M_n = \frac{1}{n} \sum_{i=1}^n X_i$. And M_n is refer to is refer to as the sample mean of (x_1, \dots, x_n)

Remarks:

1. M_n is itself a random variable
2. For large M , we expect that M_n is "close" to $E[X]$

Example 1 Let random variable x indicate the outcome of tossing a die (possibly biased). Let random variable I be defined as:

$$I = \begin{cases} 1, & \text{if } x = 5, \\ 0, & \text{otherwise.} \end{cases}$$

Let $M_n = \frac{1}{n} \sum_{i=1}^n I_i$ For a given "M".

Remarks:

1. M_n is the relative frequency (which we will also denote by f_5^n) at which "5" is seen.
2. $E[I] = 0 \cdot P[I = 0] + 1 \cdot P[I = 1]$
 $= 1 \cdot P[I = 1]$
 $= P[X = 5]$

3. We all knew that f_5^n is "close" to $P[x = 5]$ for large n .

Theorem: Weak Law of Large Number

Let x_1, \dots, x_n be iid with mean μ , then for any $\epsilon > 0$. $P[|m_n - \mu| < \epsilon] > 1 - \epsilon$ for sufficiently large n .

(Equivalently: $\lim_{n \rightarrow \infty} P[|m_n - \mu| < \epsilon] = 1$; for any $\epsilon > 0$)

Apply weak law of large number to the previous example:

$$\lim_{n \rightarrow \infty} P[|f_5^n - P[x = 5]| < \epsilon] = 1$$

Theorem: Strong Law of Large Number

Let x_1, \dots, x_n be iid with mean μ and finite variance then:

$$P[\lim_{n \rightarrow \infty} M_n = \mu] = 1 \quad \text{For } (M_1, M_2, \dots, M_n)$$

Apply Strong law of large number, the previous example we have:

$$P[\lim_{n \rightarrow \infty} f_5^n - P(X = 5)] = 1$$