Multivariate Normal Distribution

September 30, 2008

1 Random Vector

A random vector \( X = (X_1, x_2, \cdots, X_k)^T \) is a vector of random variables.

1. Discrete Case

If \( X \) takes value on a finite or countable set (or each \( X_i \) is a discrete random variable), we say \( X \) is a discrete random vector. In this case, the distribution of \( X \) is driven by the joint probability function.

\[
p(t_1, t_2, \cdots, t_k) = P(X_1 = t_1, \cdots, X_k = t_k).
\]

2. Continuous Case

In this case, the distribution of \( X \) is driven by the joint probability density function \( f(x_1, \cdots, x_k) \). The joint density function \( f \) satisfies that for any measurable set \( A \subseteq \mathbb{R}^k \).

\[
P(X \in A) = \int_A f(x_1, \cdots, x_k) dx.
\]

We can also define the joint cdf, \( F \), of \( X \)

\[
F(t_1, \cdots, t_k) = P(X_1 \leq t_1, \cdots, X_k \leq t_k) = \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_k} f(x_1, \cdots, x_k) dx_1 \cdots dx_k.
\]
It is easy to see that
\[ f(x_1, \ldots, x_k) = \frac{\partial^k}{\partial x_1 \cdots \partial x_k} F(x_1, \ldots, x_k). \]

3. Moments
\[ E(X) = (E(X_1), \ldots, E(X_k))^T \]
\[ COV(X) = E((X - EX)(X - EX)^T) = E(XX^T) - E(X)E(X)^T. \]

It can be seen that for any matrix \( A \),
\[ COV(AX) = ACOV(X)A^T. \]

The moment generating function of \( X \) is defined as (for \( t \in \mathbb{R}^k \))
\[ M_X(t) = E(e^{t^T X}). \]

2 Multivariate Normal Distribution

Suppose \( X = (X_1, \ldots, X_k) \) and \( X_i \) are i.i.d. standard normal random variables. Then it is obviously that
\[ E(X) = (0, 0, \ldots, 0), COV(X) = I_k. \]

Then for a \( n \) dimensional vector \( \mu \) and \( n \times k \) matrix \( A \)
\[ E(\mu + AX) = \mu, COV(\mu + AX) = AA^T. \]

Denote \( AA^T \) by \( \Sigma \), we have the following definition.

**Definition 1** The distribution of random vector \( AX \) is called a multivariate normal distribution with covariance matrix \( \Sigma \) and is denoted by \( N(0, \Sigma) \). And the distribution of \( \mu + AX \) is called a multivariate normal distribution with mean \( \mu \) and covariance matrix \( \Sigma \), \( N(\mu, \Sigma) \).
To make the definition valid, we need to verify that the distribution of $AX$ depend on $A$ only through $AA^T$. We can use the moment generating function to do this.

Suppose the moment generating function of $X$ is $M(t)$, we know that $M(t) = e^{t^Tt/2}$.

So the m.g.f. of $AX$ is

$$M_{AX}(t) = E(e^{t^TAX}) = M(t^TA) = e^{t^TAATt}.$$ 

This means the m.g.f. of $AX$ depend on $A$ only through $AA^T$, so the distribution of $AX$ only depends on $AA^T$.

Based on the definition, we can also calculate the joint pdf of $N(\mu, \Sigma)$ (when $\Sigma$ is full rank),

$$f(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^n (\det|\Sigma|)^{-1/2} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

where $x = (x_1, \cdots, x_n)^T$ is a $n$ dimensional vector. We can also see that if $Y$ follows $N(\mu, \Sigma)$ distribution then for any matrix $B$

$$BY \sim N(B\mu, B\Sigma B^T).$$