Multivariate Gaussian Distributions and Weighted Least Squares

In statistics, the Multivariate Gaussian Distribution is a generalization of the one-dimensional (univariate) normal distribution to higher dimensions. For calculating the probabilities of things happening, we would do integral to the Gaussian. We usually do zero-mean normalization to our data x, by changing into $\boldsymbol{X} = \frac{\boldsymbol{x}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}$(minus by x's mean and divide by standard deviation). For the new X, the normal probability density function is below:

$$\boldsymbol{p}(\boldsymbol{x}) = \frac{\mathbf{1}}{\sqrt{\mathbf{2}\boldsymbol{\pi}}}\boldsymbol{e}^{-x^2/2}$$

For $P(\boldsymbol{x},\boldsymbol{y})$**,** When $x\ and\ y$ are independent, the density function is a little different below:

$$\textbf{Independent } \boldsymbol{x} \textbf{ and } \boldsymbol{y} \quad p(x,y) = \frac{1}{2\pi\sigma_1\sigma_2}\ \boldsymbol{e}^{-(x-m_1)^2/2\sigma_1^2}\,\boldsymbol{e}^{-(y-m_2)^2/2\sigma_2^2}$$

Because $\sigma_{x,y} = 0$ (independent), the two exponents can be combined in a beautiful equation. Variables m_1 and m_2 stands for the mean.

$$-\frac{(x-m_1)^2}{2\sigma_1^2} - \frac{(y-m_2)^2}{2\sigma_2^2} = -\frac{1}{2}\begin{bmatrix} x-m_1 & y-m_2 \end{bmatrix}\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\begin{bmatrix} x-m_1 \\ y-m_2 \end{bmatrix} \quad \text{(A)}$$

The equations A can be written as $X^TQ\Lambda^{-1}Q^TX$, and further decomposing into $Y\Lambda^{-1}Y^T$ when part of $\boldsymbol{y}$ is decided by $\boldsymbol{x}$ and part is truly independent (They are correlated, so we do un-correlating). So $Y = Q^T \cdot (x-m)$.

Then, for an unsolvable system $Ax = b$, we would use weighted least squares when the errors in b are not independent or their variances are not equal. We need to minimize the error $Error = \sum_{i=1}^{m}\frac{(b-Ax)_i}{\sigma_i}\cdot\frac{(b-Ax)_i}{\sigma_i} = (b-Ax)^TV^{-1}(b-Ax)$.

Because $\frac{1}{{\sigma_i}^2}$ comes from V^{-1}, we multiply $V^{-1/2}$ on both sides and get $A^TV^{-1}b = A^TV^{-1}A\hat{x}$.

One of the special Joint Probability Distribution is called **standard bivariate normal distribution**. In this case, we have:

$$P(\boldsymbol{x_1},\boldsymbol{x_2}) = \frac{\mathbf{1}}{\mathbf{2}\,\boldsymbol{\pi}\sqrt{\mathbf{1}-\boldsymbol{\rho}^{\mathbf{2}}}}\cdot\boldsymbol{e}^{\frac{-\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}+\frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{-\sigma_1\sigma_2}\right)}{2(1-\rho^2)}}$$

In the equation, $\boldsymbol{\rho}$ is the correlation coefficient which equals to $\frac{\boldsymbol{V_{1,2}}}{\boldsymbol{\sigma_1\sigma_2}}(V_{1,2}\ is\ the\ covarience)$.

By the way, $(b-Ax)^T\left(V^{-1/2}\right)^TV^{-1/2}(b-Ax)$, we can compute a covariance matrix A, and use it to generate the random numbers which are corresponding to Normal Distribution.