Multivariate Probability Review

Given a "random vector", $\underset{\rightarrow}{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}$,

A "center" of the dist'n is the mean vector, $\underset{\rightarrow}{\mu} = E\underset{\rightarrow}{X} = \begin{pmatrix} EX_1 \\ \vdots \\ EX_d \end{pmatrix}$

A "measure of spread" is the covariance matrix:

$$\Sigma = \operatorname{cov}(X) = \begin{pmatrix} \operatorname{var}(X_1) & \cdots & \operatorname{cov}(X_1, X_d) \\ \vdots & \ddots & \vdots \\ \operatorname{cov}(X_d, X_1) & \cdots & \operatorname{var}(X_d) \end{pmatrix}$$

Multivariate Probability Review, (cont.)

Covariance matrix:

- Nonegative Definite (since all variances are ≥ 0)
- Provides "elliptical summary of distribution"
- Calculated via "outer product":

$$\Sigma = \operatorname{cov}(X) = E\begin{pmatrix} (X_1-\mu_1)(X_1-\mu_1) & \cdots & (X_1-\mu_1)(X_d-\mu_d) \\ \vdots & \ddots & \vdots \\ (X_d-\mu_d)(X_1-\mu_1) & \cdots & (X_d-\mu_d)(X_d-\mu_d) \end{pmatrix} =$$

$$\Sigma = E\left(\underset{\rightarrow}{X} - \underset{\rightarrow}{\mu}\right)\left(\underset{\rightarrow}{X} - \underset{\rightarrow}{\mu}\right)^t$$

Multivariate Probability Review, (cont.)

Empirical versions:

Given a "random sample" $\underset{\rightarrow}{X}_1, \ldots, \underset{\rightarrow}{X}_n$,

Estimate the "theoretical mean" $\underset{\rightarrow}{\mu}$, with the "sample mean":

$$\hat{\underset{\rightarrow}{\mu}} = \underset{\rightarrow}{\overline{X}} = \begin{pmatrix} \overline{X}_1 \\ \vdots \\ \overline{X}_d \end{pmatrix} = \frac{1}{n}\sum_{i=1}^{n} \underset{\rightarrow}{X}_i$$

Multivariate Probability Review, (cont.)

Empirical versions (cont.)

And estimate the “theoretical cov.”, with the “sample cov.”:

$$\hat{\Sigma} = \frac{1}{n-1} \begin{pmatrix} \sum_{i=1}^{n} (X_{i1} - \overline{X}_1)^2 & \cdots & \sum_{i=1}^{n} (X_{i1} - \overline{X}_1)(X_{id} - \overline{X}_d) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} (X_{id} - \overline{X}_d)(X_{i1} - \overline{X}_1) & \cdots & \sum_{i=1}^{n} (X_{id} - \overline{X}_d)^2 \end{pmatrix}$$

Normalizations:

- $\frac{1}{n-1}$ gives unbiasedness
- $\frac{1}{n}$ gives MLE in Gaussian case

Multivariate Probability Review, (cont.)

Outer product representation:

$$\hat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}\begin{pmatrix} (X_{i1}-\overline{X}_1)^2 & \cdots & (X_{i1}-\overline{X}_1)(X_{id}-\overline{X}_d) \\ \vdots & \ddots & \vdots \\ (X_{id}-\overline{X}_d)(X_{i1}-\overline{X}_1) & \cdots & (X_{id}-\overline{X}_d)^2 \end{pmatrix}$$

$$\hat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\underrightarrow{X_i} - \underrightarrow{\overline{X}}\right)\left(\underrightarrow{X_i} - \underrightarrow{\overline{X}}\right)^t = \tilde{X}\tilde{X}^t,$$

where: $$\tilde{X} = \frac{1}{\sqrt{n-1}}\left(\underline{X_1} - \underline{\overline{X}} \quad \cdots \quad \underline{X_n} - \underline{\overline{X}}\right)_{d\times n}$$

[go to next part]