ECSE 6520: Estimation and Detection Theory

Multivariate Gaussian Distribution

Class Notes - 3

January 26, 2014

Contents

1 The Multivariate Gaussian Distribution

- A key concept in statistical inference is that of the *statistical model* which is simply a hypothesized probability distribution or density function $f(x)$ for the observed data.
- Broadly stated statistical inference explores the possibility of fitting a given model to the data x.
- To simplify this task it is common to restrict $f(x)$ to a class of parametric models $\{f_\theta(x)\}_{\theta\in\Theta}$, where $f_\theta(x)$ is a known function and θ is a vector of unknown parameters taking values in a parameter space Θ.
- In this case statistical inference boils down to inferring properties of the true value of θ parameterizing $f_\theta(x)$ that generated the data sample x.

The Gaussian distribution play a major role in parametric statistical inference and is widely employed in statistical signal processing. Some reasons for this include:

- Relative simplicity and tractability.
- Estimators and detectors with intuitive forms and properties.
- Justification in terms of Central Limit Theorem.

2 Characteristic Function

Definition 1 *Characteristic Function - The characteristic function of an N dimensional random variable X is defined as*

$$\Phi(\omega) = E[e^{j\omega^T X}] = \int e^{j\omega^T x} f(x) dx.$$

- The characteristic function of a random variables uniquely characterizes the random variable.

- The characteristic function of a multivariate Gaussian random variable, $X \sim \mathcal{N}(\mu, \Sigma)$, is given by

$$\Phi(\omega) = e^{-j\omega\mu - \frac{1}{2}\omega^T\Sigma\omega}.$$

3 Useful Facts about Multivariate Gaussian Distribution

- **Unimodality and symmetry of the Gaussian density:** The multivariate Gaussian density is unimodal (has a unique maximum) and is symmetric about its mean parameter.

- **Uncorrelated Gaussian random variables are independent:** When the covariance matrix Σ is diagonal, i.e., $cov(Xi, Xj) = 0$, for $i \neq j$, then the multivariate Gaussian density reduces to a product of univariate densities

$$f(X) = \prod_{i=1}^{n} f(Xi).$$

- **Marginals of a multivariate Gaussian density are Gaussian:** If $X = [X_1, ..., X_n]^T$ is multivariate Gaussian then any subset of the elements of X is also Gaussian. In particular X_1 is univariate Gaussian and $[X_1, X_2]$ is bivariate Gaussian.

- **Linear combinations are Gaussian:** Let $X = [X_1, ..., X_n]^T$ be a multivariate Gaussian random vector and let H be a $p \times n$ non-random matrix. Then $Y = HX$ is a vector of linear combinations of the X_i's. The distribution of Y is multivariate (p-variate) Gaussian with mean $E[Y] = HE[X]$ and $p \times p$ covariance matrix $cov(Y) = Hcov(X)H^T$.

- **The conditional distribution of a Gaussian given another Gaussian is Gaussian:** Let the vector $X = [X_1, ..., X_p]^T$ and $Y = [Y_1, ..., Y_q]^T$ be p-variate and q-variate

Gaussian random variables, respectively. Let the mean value and covariance matrix of X and Y be μ_X and μ_Y and Σ_X and Σ_Y, respectfully. Then the conditional density $f_{Y|X}(y|x)$ of Y given $X = x$ is multivariate (q-variate) Gaussian. The conditional mean, $\mu_{Y|X}$, is given by

$$\mu_{Y|X} = E[Y|X = x] = \mu_Y + \Sigma_{X,Y}^T \Sigma_X^{-1}(x - \mu_X)$$

where $\Sigma_{X,Y}^T = E[(X - \mu_X)(Y - \mu_Y)^T]$.

The conditional covariance, $cov(Y|X = x)$, is given by

$$cov(Y|X = x) = E[(Y - \mu_{Y|X})(Y - \mu_{Y|X})^T|X = x] = \Sigma_Y - \Sigma_{X,Y}^T \Sigma_X^{-1} \Sigma_{X,Y}^T.$$

4 Central Limit Theorem

Theorem 1 *Let X_i, $i = 1, ..., n$ be independent identically distributed random vectors in $\mathbb{R}^p$ with common mean $E[X_i] = \mu$ and finite positive definite covariance matrix $cov(X_i) = \Sigma$. Then as n goes to infinity the distribution of the random vector*

$$Z_n = \sum_{i=1}^{n} \frac{(X_i - \mu)}{\sqrt{n}}$$

converges to a p-variate Gaussian distribution with zero mean and covariance Σ.