Tutorial 2: Maximum-Likelihood and Bayesian Parameter Estimation

Rui Zhao

rzhao@ee.cuhk.edu.hk

Jan. 23, 2014

Outline

1. ML estimate: $\hat{\Sigma}$ of multivariate Gaussian
2. ML estimate bias: $\hat{\sigma}^2$ of univariate Gaussian
3. Bayesian estimate: Univariate Gaussian
   - posteriori $p(\mu|\mathcal{D})$
   - conditional probability density $p(x|w_i, \mathcal{D}_i)$

Outline

1. ML estimate: $\hat{\Sigma}$ of multivariate Gaussian
2. ML estimate bias: $\hat{\sigma}^2$ of univariate Gaussian
3. Bayesian estimate: Univariate Gaussian
   - posteriori $p(\mu|\mathcal{D})$
   - conditional probability density $p(x|w_i, \mathcal{D}_i)$

Outline

1. ML estimate: $\hat{\Sigma}$ of multivariate Gaussian
2. ML estimate bias: $\hat{\sigma}^2$ of univariate Gaussian
3. Bayesian estimate: Univariate Gaussian
   - posteriori $p(\mu|\mathcal{D})$
   - conditional probability density $p(x|w_i, \mathcal{D}_i)$

1. ML estimate: $\hat{\Sigma}$ of multivariate Gaussian

Multivariate Gassian Case: unknown μ and Σ

$$\hat{\mu} = \frac{1}{n}\sum_{k=1}^{n}\mathbf{x}_k \quad (1)$$

$$\hat{\mathbf{\Sigma}} = \frac{1}{n}\sum_{k=1}^{n}(\mathbf{x}_k - \hat{\mu})(\mathbf{x}_k - \hat{\mu})^t \quad (2)$$

- $\hat{\mu}$ is the sample mean.
- $\hat{\mathbf{\Sigma}}$ is the arithmetic average of the n matrices $(\mathbf{x}_k - \hat{\mu})(\mathbf{x}_k - \hat{\mu})^t$.

1. ML estimate: $\hat{\Sigma}$ of multivariate Gaussian

Multivariate normal density

$$p(\mathbf{x} \mid \mu,\ \mathbf{\Sigma}) = \frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mu)^t \mathbf{\Sigma}^{-1}(\mathbf{x}-\mu) \right] \quad (3)$$

Draw $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ independently from $p(\mathbf{x} \mid \mu,\ \mathbf{\Sigma})$, and the joint density (likelihood) is:

$$p(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n \mid \mu, \mathbf{\Sigma}) = \quad (4)$$

$$\frac{1}{(2\pi)^{nd/2}|\mathbf{\Sigma}|^{n/2}} \exp\left[-\frac{1}{2} \sum_{k=1}^{n} (\mathbf{x}_k-\mu)^t \mathbf{\Sigma}^{-1}(\mathbf{x}_k-\mu) \right] \quad (5)$$

Log-likelihood $l(\mu,\ \mathbf{\Sigma})$ is

$$l(\mu, \mathbf{\Sigma}) = -\frac{nd}{2}\ln(2\pi) - \frac{n}{2}\ln|\mathbf{\Sigma}| - \frac{1}{2}\sum_{k=1}^{n}(\mathbf{x}_k-\mu)^t \mathbf{\Sigma}^{-1}(\mathbf{x}_k-\mu) \quad (6)$$

1. ML estimate: $\hat{\Sigma}$ of multivariate Gaussian

Multivariate normal density

Let $A = \mathbf{\Sigma}^{-1}$

$$l(\mu, \mathbf{\Sigma}) = -\frac{nd}{2}\ln(2\pi) - \frac{n}{2}\ln|\mathbf{\Sigma}| - \frac{1}{2}\sum_{k=1}^{n}(\mathbf{x}_k - \mu)^t \mathbf{\Sigma}^{-1}(\mathbf{x}_k - \mu) \quad (7)$$

$$l(\mu, \mathbf{\Sigma}) = -\frac{nd}{2}\ln(2\pi) + \frac{n}{2}\ln\mathbf{A} - \frac{1}{2}\sum_{k=1}^{n}(\mathbf{x}_k - \mu)^t \mathbf{A}(\mathbf{x}_k - \mu) \quad (8)$$

$$\frac{\partial l(\mu, \mathbf{\Sigma})}{\partial \mathbf{A}} = \frac{n}{2}\mathbf{A}^{-1} - \frac{1}{2}\sum_{k=1}^{n}(\mathbf{x}_k - \mu)(\mathbf{x}_k - \mu)^t = 0 \quad (9)$$

Replace $\mathbf{A}$ by $\mathbf{\Sigma}^{-1}$

$$\hat{\mathbf{\Sigma}} = \frac{1}{n}\sum_{k=1}^{n}(\mathbf{x}_k - \hat{\mu})(\mathbf{x}_k - \hat{\mu})^t, \quad (\hat{\mu} = \frac{1}{n}\sum_{k=1}^{n}\mathbf{x}_k) \quad (10)$$

Outline

1. ML estimate: $\hat{\Sigma}$ of multivariate Gaussian
2. ML estimate bias: $\hat{\sigma}^2$ of univariate Gaussian
3. Bayesian estimate: Univariate Gaussian
    - posteriori $p(\mu|\mathcal{D})$
    - conditional probability density $p(x|w_i, \mathcal{D}_i)$

Outline

1. ML estimate: $\hat{\Sigma}$ of multivariate Gaussian
2. ML estimate bias: $\hat{\sigma}^2$ of univariate Gaussian
3. Bayesian estimate: Univariate Gaussian
   - posteriori $p(\mu|\mathcal{D})$
   - conditional probability density $p(x|w_i, \mathcal{D}_i)$