Lecture Outline

Reading: Chapter 1

- Gaussian Distribution
- Estimation of Distributional Parameters from Data
- Curve Fitting Reviewed

1.2 Probability Theory

1.2.4 Gaussian Distribution

The *Gaussian* or *normal* distribution is defined as

$$\begin{aligned} p(x) &\equiv \mathcal{N}(x|\mu,\sigma^2) \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} \end{aligned} \tag{1}$$

where the two parameters are mean μ and variance σ^2. Figure 1 shows the Gaussian pdf.

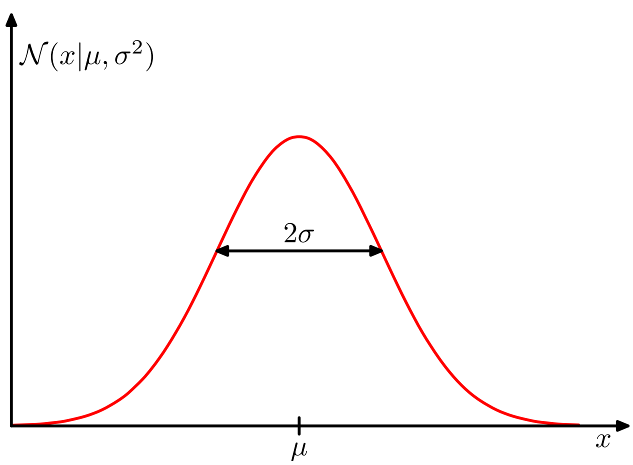

Figure 1: (Figure 1-13)

We have the following:

- The *precision* is $\beta = 1/\sigma^2$
- The *mode* is the maximum of the distribution. For a Gaussian distribution the mode coincides with the mean.
- The *first moment* or *mean* is calculated as

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} p(x)x dx = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu,\sigma^2)x dx = \mu \tag{2}$$

©2016 Phillip De Leon

- The *second moment* is calculated as

$$\mathbb{E}[x^2] = \int_{-\infty}^{\infty} p(x)x^2 dx = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu,\sigma^2)x^2 dx = \mu^2 + \sigma^2 \tag{3}$$

- The *variance* is calculated as

$$\mathrm{var}[x] = \sigma^2 = \mathbb{E}[x^2] - \mathbb{E}[x]^2. \tag{4}$$

For a normally-distributed x, the probability of x falling within one standard deviation of the mean is given by

$$p(x \in (\mu-\sigma, \mu+\sigma)) = \int_{\mu-\sigma}^{\mu+\sigma} \mathcal{N}(x|\mu,\sigma^2)dx = 0.682 \tag{5}$$

The D-dimensional, Gaussian random vector $\mathbf{x}$ is given by

$$\mathcal{N}(x|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad = \quad \frac{1}{(2\pi)^{D/2}|\boldsymbol{\Sigma}|^{1/2})} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\} \tag{6}$$

where $\boldsymbol{\mu}$ is the D-dimensional mean vector, $\boldsymbol{\Sigma}$ is the $D \times D$ covariance matrix,

$$\boldsymbol{\Sigma} \quad = \quad \mathbb{E}\left[\mathbf{x}\mathbf{x}^T\right] - \mathbb{E}\left[\mathbf{x}\right]\mathbb{E}\left[\mathbf{x}^T\right] \tag{7}$$

and $|\boldsymbol{\Sigma}|$ is the matrix determinant of $\boldsymbol{\Sigma}$.

Estimation of Distributional Parameters from Data

In this course, we will often want to statistically model data using a distribution and will need to estimate the distribution's parameters. Suppose that we have a data set of scalar observations

$$\mathsf{x} \quad = \quad (x_1, \ldots, x_N)^T \tag{8}$$

drawn independently from a Gaussian distribution, $\mathcal{N}(x|\mu,\sigma^2)$, i.e. independent and identically distributed (i.i.d.). Because our data is i.i.d., the likelihood of the data set given model parameters μ, σ^2 is the product of the likelihoods of each data point

$$p(\mathsf{x}|\mu,\sigma^2) \quad = \quad \prod_{n=1}^{N} \mathcal{N}(x_n|\mu,\sigma^2) \tag{9}$$

as shown in Fig. 2.

One method for determining the distribution parameters from data is to maximize the likelihood function or equivalently, maximize the log-likelihood (for better numerical precision)

$$\ln p(\mathsf{x}|\mu,\sigma^2) \quad = \quad \sum_{n=1}^{N} \ln \mathcal{N}(x_i|\mu,\sigma^2) \tag{10}$$

Maximizing the likelihood leads leads to the following solutions for Gaussian data:

$$\mu_{\mathrm{ML}} \quad = \quad \frac{1}{N}\sum_{n=1}^{N} x_n \tag{11}$$

and

$$\sigma^2_{\mathrm{ML}} \quad = \quad \frac{1}{N}\sum_{n=1}^{N}(x_n - \mu_{\mathrm{ML}})^2 \tag{12}$$

©2016 Phillip De Leon

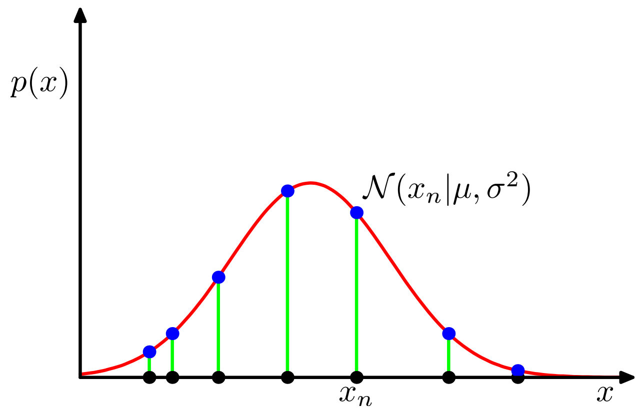

Figure 2: (Figure 1-14–see caption in text!)

which are the sample mean and sample variance respectively. The ML approach on average yields the correct mean

$$\mathbb{E}[\mu_{\mathrm{ML}}] = \mu \tag{13}$$

but underestimates the variance

$$\mathbb{E}[\sigma^2_{\mathrm{ML}}] = \left(\frac{N-1}{N}\right)\sigma^2. \tag{14}$$

However, as the number of data points increases the bias decreases. This effect is shown below

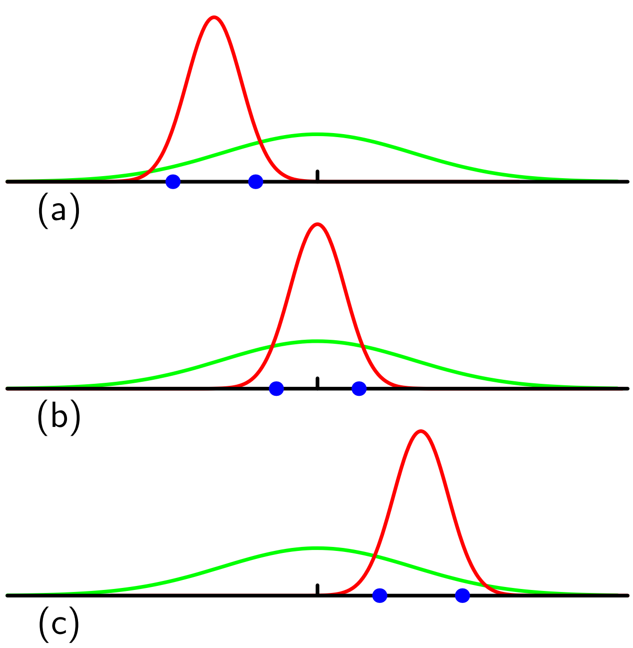

Figure 3: (Figure 1-15)

Note: Some pdfs, such as a Gaussian Mixture Model (GMM), do not have closed-form ML solutions for distribution parameters. In these cases, we turn to iterative approaches to estimate parameters.