Homework 2 Solutions

The solutions of Exercises 3.2 – 3.5 are on Murphy's website.

Problem 1:

a.) $\sigma = 1,\ \mu = 0$: $\quad P(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)$

$$\Pr(-a \le x \le a) = P(a) - P(-a) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{a}{\sqrt{2}}\right) - \mathrm{erf}\left(-\frac{a}{\sqrt{2}}\right)\right] = \mathrm{erf}\left(\frac{a}{\sqrt{2}}\right)$$

in Matlab, Python etc we find $\mathrm{erf}\left(\frac{a}{\sqrt{2}}\right) = \mathrm{erf}\left(\frac{1.96}{\sqrt{2}}\right) = 0.95$ ✓

b.) $P_{\sigma\mu}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt$

confidence interval is centered at μ so write $[\mu-\delta, \mu+\delta]$. Find δ:

$$0.95 = P_{\sigma\mu}(\mu+\delta) - P_{\sigma\mu}(\mu-\delta)$$

$$= \int_{\mu-\delta}^{\mu+\delta} \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt$$

subst: $s = \frac{t-\mu}{\sigma}$

$ds = \frac{dt}{\sigma}$

$t = \mu \pm \delta \Rightarrow s = \pm\frac{\delta}{\sigma}$

$$= \int_{-\delta/\sigma}^{\delta/\sigma} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}s^2\right) ds$$

$$= \frac{1}{2}\left[\mathrm{erf}\left(\frac{\delta}{\sqrt{2}\,\sigma}\right) - \mathrm{erf}\left(-\frac{\delta}{\sqrt{2}\,\sigma}\right)\right]$$

$$= \mathrm{erf}\left(\frac{\delta}{\sqrt{2}\,\sigma}\right) \quad \Rightarrow \quad \frac{\delta}{\sigma} = a = 1.96 \quad \Rightarrow \quad \delta = 1.96 \cdot \sigma$$

Answer: 95% confidence interval is $[\mu - 1.96\cdot\sigma,\ \mu + 1.96\cdot\sigma]$

HW 2 p2

Problem 2 $f(x) = x^T A x + b^T x + c$

$$= \sum_{ij} a_{ij} x_i x_j + \sum_j b_j x_j + c$$

Note: $\frac{\partial}{\partial x_k} x_i = \delta_{ik}$ (Kronecker-delta: $\delta_{ik} = \begin{cases} 1 & \text{if } i=k \\ 0 & \text{if } i \neq k \end{cases}$)

and by the product rule: $\frac{\partial}{\partial x_k}(x_i \cdot x_j) = x_i \cdot \delta_{kj} + \delta_{ki} x_j$

Thus: $\frac{\partial}{\partial x_k} f(x) = \sum_{ij} a_{ij} \frac{\partial}{\partial x_k}(x_i x_j) + \sum_j b_j \frac{\partial}{\partial x_k} x_j$

$$= \sum_{ij} a_{ij} x_i \cdot \delta_{kj} + \sum_{ij} a_{ij} x_j \delta_{ki} + \sum_j b_j \delta_{kj}$$

$\downarrow$ j-sum goes away $j=k$

$$= \sum_i a_{ik} x_i + \sum_j a_{kj} x_j + b_k$$

$$= \sum_i (A^T)_{ki} x_i + \sum_j a_{kj} x_j + b_k$$

$$= [A^T x]_k + [Ax]_k + [b]_k \quad \Rightarrow \quad \nabla f(x) = (A^T + A)x + b$$

we had

$$NLL(\mu, \Sigma) = -\frac{N}{2} \log |\Sigma^{-1}| + \frac{1}{2} \sum_n (y_n - \mu) \Sigma^{-1} (y_n - \mu) \quad \Leftarrow (4.39)$$

$$\Rightarrow \frac{\partial}{\partial \mu_k} NLL(\mu, \Sigma) = \frac{1}{2} \sum_{n=1}^{N} (\Sigma^{-T} + \Sigma^{-1})(y_n - \mu)$$

$\Sigma^{-T} = \Sigma^{-1}$ since Σ is symmetric!

$$= \Sigma^{-1} \cdot \sum_{n=1}^{N} (y_n - \mu) \stackrel{!}{=} 0$$

Σ^{-1} is invertible

$$\sum_{n=1}^{N} (y_n - \mu) = 0 \Rightarrow \sum_{n=1}^{N} y_n = N \cdot \mu$$

$$\Rightarrow \mu = \frac{1}{N} \sum_{n=1}^{N} y_n$$