

# PRML Solutions

## C01 Introduction

Zhao Yang

Department of Automation, Tsinghua University

**Ex 1.1(\*).** *Solutions.* We have  $y(x, \mathbf{w}) = \sum_{j=0}^M w_j x^j$  and  $E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2$  then differentiate with respect to  $w_i$ , we obtain

$$\frac{\partial E}{\partial w_i} = \sum_{n=1}^N (y(x_n, \mathbf{w}) - t_n) x_n^i = 0 \quad (1)$$

which is equal to  $\sum_{n=1}^N \sum_{j=0}^M w_j x_n^{j+i} = \sum_{n=1}^N t_n x_n^i$ .  
Let  $A_{ij} = \sum_{n=1}^N x_n^{i+j}$  and  $T_i = \sum_{n=1}^N x_n^i t_n$ , so we get

$$\sum_{j=0}^M A_{ij} w_j = T_i$$

□

**Ex 1.2(\*).** *Solutions.* We have  $E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2} \sum_{i=0}^M w_i^2$  then differentiate with respect to  $w_i$ , we obtain

$$\frac{\partial E}{\partial w_i} = \sum_{n=1}^N (y(x_n, \mathbf{w}) - t_n) x_n^i + \lambda w_i = 0 \quad (2)$$

which is equal to  $\sum_{n=1}^N \sum_{j=0}^M w_j x_n^{j+i} + \lambda w_i = \sum_{n=1}^N t_n x_n^i$ .  
Let  $A_{ij} = \sum_{n=1}^N x_n^{i+j}$  and  $T_i = \sum_{n=1}^N x_n^i t_n$ , we have

$$\sum_{j=0}^M A_{ij} w_j + \lambda w_i = T_i$$

Let  $\tilde{A}_{ij} = A_{ij} + \lambda I_{ij}$ , so we get

$$\sum_{j=0}^M \tilde{A}_{ij} w_j = T_i$$

□

**Ex 1.3(\*\*).** *Solutions.* Let us denote apples, oranges and limes by  $a$ ,  $o$  and  $l$  respectively.

$$\begin{aligned} p(a) &= p(a|r)p(r) + p(a|b)p(b) + p(a|g)p(g) \\ &= 0.3 \times 0.2 + 0.5 \times 0.2 + 0.3 \times 0.6 \\ &= 0.34 \end{aligned} \tag{3}$$

and by using Bayes' theorem, we obtain

$$\begin{aligned} p(g|o) &= \frac{p(o|g)p(g)}{\sum_{x=r,g,b} p(o|x)p(x)} \\ &= 0.5 \end{aligned} \tag{4}$$

□

**Ex 1.4(\*\*).** *Solutions.*

□

**Ex 1.5(\*).** *Proof.*

$$\begin{aligned} \mathbb{E}[(f(x) - \mathbb{E}f(x))^2] &= \mathbb{E}[f^2(x) - 2f(x)\mathbb{E}f(x) + \mathbb{E}^2f(x)] \\ &= \mathbb{E}f^2(x) - \mathbb{E}^2f(x) \end{aligned}$$

□

**Ex 1.6(\*).** *Proof.* if  $x$  and  $y$  are independent, we have  $p(x, y) = p(x)p(y)$ . So we obtain

$$\begin{aligned} Cov[x, y] &= \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y] \\ &= \iint xyp(x, y)dxdy - \int xp(x)dx \int yp(y)dy \\ &= \iint xyp(x)p(y)dxdy - \iint xyp(x)p(y)dxdy \\ &= 0 \end{aligned}$$

□

**Ex 1.7(\*\*).** *Proof.* From Cartesian to polar coordinates, the transformation is defined by

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned}$$

The Jacobian of the change of variables is given by

$$\begin{aligned} \frac{\partial(x, y)}{\partial(r, \theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} \\ &= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} \\ &= r \end{aligned}$$

So, we have

$$\begin{aligned}
I^2 &= \int_0^{2\pi} \int_0^\infty e^{-\frac{r^2}{2\sigma^2}} r dr d\theta \\
&= 2\pi \int_0^\infty e^{-\frac{u}{2\sigma^2}} \frac{1}{2} du \\
&= 2\pi\sigma^2
\end{aligned}$$

Thus

$$I = (2\pi\sigma^2)^{1/2} \quad (5)$$

Finally, we obtain

$$\begin{aligned}
y &= \int_{-\infty}^{+\infty} \mathcal{N}(x|\mu, \sigma^2) dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{1}{2\pi\sigma^2} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2\sigma^2}} dt \\
&= \frac{I^2}{2\pi\sigma^2} \\
&= 1
\end{aligned}$$

□

**Ex 1.8(\*\*).** *Proof.* We have

$$\begin{aligned}
\mathbb{E}[x] &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} x dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \left( \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} (x - \mu) dx + \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mu dx \right) \\
&= \frac{\mu}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \mu
\end{aligned} \quad (6)$$

Because we have  $\int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} (x - \mu)^2 dx = \sqrt{2\pi\sigma^2}$ , then we differentiate both side of uppon with respect to  $\sigma^2$ . We obtain

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} (x - \mu)^2 dx = \sigma^2 \quad (7)$$

Thus

$$\begin{aligned}
\mathbb{E}[x^2] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} [(x - \mu)^2 + 2x\mu - \mu^2] dx \\
&= \sigma^2 + 2\mu^2 - \mu^2 \\
&= \sigma^2 + \mu^2
\end{aligned}$$

□

**Ex 1.9(\*).** *Solutions.* We simply differentiate it with respect to  $x$  to obtain

$$\frac{d}{dx}\mathcal{N}(x|\mu, \sigma^2) = -\mathcal{N}(x|\mu, \sigma^2)\frac{x - \mu}{\sigma^2} = 0 \quad (8)$$

Thus  $x = \mu$ . Similarly, we obtain

$$\frac{\partial}{\partial \mathbf{x}}\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \quad (9)$$

Thus  $\mathbf{x} = \boldsymbol{\mu}$  □

**Ex 1.10(\*).** *Solutions.* Because of the independence, we have  $p(x, z) = p(x)p(z)$ . Thus

$$\begin{aligned} \mathbb{E}[x + z] &= \iint (x + z)p(x, z)dx dz \\ &= \iint (x + z)p(x)p(z)dx dz \\ &= \int xp(x)dx + \int zp(z)dz \\ &= \mathbb{E}[x] + \mathbb{E}[z] \end{aligned}$$

Similarly, for variance, we have

$$\begin{aligned} Var[x + z] &= \mathbb{E}[\{x + z - \mathbb{E}[x + z]\}^2] \\ &= \iint \{x + z - \mathbb{E}[x + z]\}^2 p(x)p(z)dx dz \\ &= \iint \{x + z - \mathbb{E}[x] - \mathbb{E}[z]\}^2 p(x)p(z)dx dz \\ &= \int (x - \mathbb{E}x)^2 p(x)dx + \int (z - \mathbb{E}z)^2 p(z)dz + \\ &\quad 2 \iint (x - \mathbb{E}x)(z - \mathbb{E}z)p(x)p(z)dx dz \\ &= Var[x] + Var[z] \end{aligned}$$

□

**Ex 1.11(\*).** *Solutions.* Too trivial to do it. □

**Ex 1.12(\*\*).** *Solutions.* If  $n = m$ , we obtain  $\mathbb{E}[x_n x_m] = \mathbb{E}[x_n^2] = \mu^2 + \sigma^2$ . If  $n \neq m$ , we obtain  $\mathbb{E}[x_n]\mathbb{E}[x_m] = \mu^2$ . Thus

$$\mathbb{E}[\mu_{ML}] = \frac{1}{N} \sum_{n=1}^N \mathbb{E}[x_n] = \mu \quad (10)$$

and

$$\begin{aligned}
\mathbb{E}[\sigma_{ML}^2] &= \frac{1}{N} \sum_{n=1}^N \mathbb{E}[(x_n - \mu_{ML})^2] \\
&= \frac{1}{N} \sum_{n=1}^N \mathbb{E}[x_n^2 - 2\frac{x_n}{N} \sum_{i=1}^N x_i + \frac{1}{N^2} (\sum_{i=1}^N x_i)^2] \\
&= \left(\frac{N-1}{N}\right) \sigma^2
\end{aligned} \tag{11}$$

□

**Ex 1.13(\*).** *Proof.*

$$\begin{aligned}
\sigma_{ML}^2 &= \frac{1}{N} \sum_{n=1}^N (x_n - \mu)^2 \\
\mathbb{E}[\sigma_{ML}^2] &= \frac{1}{N} \sum_{n=1}^N \mathbb{E}[x_n^2 - 2\mu x_n + \mu^2] \\
&= \frac{1}{N} \sum_{n=1}^N (\mu^2 + \sigma^2 - 2\mu^2 + \mu^2) \\
&= \sigma^2
\end{aligned}$$

□

**Ex 1.14(\*).** *Proof.* We want to rewrite the  $\mathbf{w}$  as  $\mathbf{w} = \mathbf{w}^S + \mathbf{w}^A$  where  $\mathbf{w}^S$  satisfy  $\mathbf{w}^S = (\mathbf{w}^S)^T$  and  $\mathbf{w}^A = -(\mathbf{w}^A)^T$ . So we have

$$\begin{aligned}
\mathbf{w}^T &= (\mathbf{w}^S)^T + (\mathbf{w}^A)^T \\
\mathbf{w} &= \mathbf{w}^S + \mathbf{w}^A
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbf{w}^S &= \frac{\mathbf{w} + \mathbf{w}^T}{2} \\
\mathbf{w}^A &= \frac{\mathbf{w} - \mathbf{w}^T}{2}
\end{aligned}$$

So

$$\begin{aligned}
\sum_{i=1}^D \sum_{j=1}^D w_{ij}^S x_i x_j &= \sum_{i=1}^D \sum_{j=1}^D \frac{w_{ij} + w_{ji}}{2} x_i x_j \\
&= \frac{1}{2} \sum_{i=1}^D \sum_{j=1}^D w_{ij} x_i x_j + \frac{1}{2} \sum_{i=1}^D \sum_{j=1}^D w_{ji} x_i x_j \\
&= \sum_{i=1}^D \sum_{j=1}^D w_{ij} x_i x_j
\end{aligned}$$

Thus, the number of independent parameters in the matrix  $w_{ij}^S$  is

$$num(i.d) = \frac{D^2 - D}{2} + D = \frac{D(D+1)}{2} \quad (12)$$

□

**Ex 1.15(\*\*\*)**. *Proof*.

$$\begin{aligned} n(D, M) &= \sum_{i_1=1}^D \sum_{i_2=1}^{i_1} \cdots \sum_{i_M=1}^{i_{M-1}} 1 \\ &= \sum_{i_1=1}^D \left\{ \sum_{i_2=1}^{i_1} \cdots \sum_{i_M=1}^{i_{M-1}} 1 \right\} \\ &= \sum_{i_1=1}^D n(i_1, M-1) \end{aligned}$$

Then use proof by induction to show that the following result holds

$$\sum_{i=1}^D \frac{(i+M-2)!}{(i-1)!(M-1)!} = \frac{(D+M-1)!}{(D-1)!M!}$$

If  $D = 1$ , it is simple to show that the result holds.

If  $D = k$ , we assume the result holds.

If  $D = k+1$ , we obtain

$$\begin{aligned} \sum_{i=1}^{k+1} \frac{(i+M-2)!}{(i-1)!(M-1)!} &= \sum_{i=1}^k \frac{(i+M-2)!}{(i-1)!(M-1)!} + \frac{(k+M-1)!}{k!(M-1)!} \\ &= \frac{(k+M-1)!}{(k-1)!M!} + \frac{(k+M-1)!}{k!(M-1)!} \\ &= \frac{(k+M)!}{k!M!} \end{aligned}$$

so the equation is proved.

Finally, we want to prove the following result holds

$$n(D, M) = \frac{(D+M-1)!}{(D-1)!M!}$$

If  $M = 2$ , we have  $n(D, 2) = \frac{D(D+1)}{2}$  which is proved in equation 12.

We assume that when  $M = k$ , the equation holds. If  $M = k+1$ , we obtain

$$\begin{aligned} n(D, k+1) &= \sum_{i=1}^D n(i, k) \\ &= \sum_{i=1}^D \frac{(i+M-1)!}{(i-1)!M!} \\ &= \frac{(D+M)!}{(D-1)!(M+1)!} \end{aligned}$$

□

**Ex 1.16(\*\*).** *Proof.* If  $M = 0$ , for all  $D$ , we have  $N(D, 0) = 1$ , which is obviously true. We assume that when  $M = k$ , the equation holds.

If  $M = k + 1$ , we have

$$\begin{aligned}
 N(D, k + 1) &= \sum_{m=0}^k n(D, m) + n(D, k + 1) \\
 &= N(D, k) + n(D, k + 1) \\
 &= \frac{(D + k)!}{D!k!} + \frac{(D + k)!}{(D - 1)!(k + 1)!} \\
 &= \frac{(D + k + 1)!}{D!(k + 1)!}
 \end{aligned}$$

Thus, if  $D \gg M$ , we have

$$\begin{aligned}
 N(D, M) &= \frac{(D + M)!}{D!M!} \\
 &\simeq \frac{(D + M)^{D+M} e^{-(D+M)}}{D^D e^{-D} M!} \\
 &= D^M \left(1 + \frac{M}{D}\right)^{D+M} \frac{e^{-M}}{M!} \\
 &\simeq D^M \left(1 + M + \frac{M^2}{D}\right) \frac{e^{-M}}{M!}
 \end{aligned}$$

which grows like  $D^M$

□

**Ex 1.17(\*\*).** *Proof.* First, we prove  $\Gamma(x + 1) = x\Gamma(x)$ .

$$\begin{aligned}
 \Gamma(x + 1) &= \int_0^\infty u^x e^{-u} du \\
 &= -u^x e^{-u} \Big|_0^\infty + \int_0^\infty x u^{x-1} e^{-u} du \\
 &= 0 + x\Gamma(x)
 \end{aligned}$$

Second, we obtain

$$\Gamma(1) = \int_0^\infty u e^{-u} du = -e^{-u} \Big|_0^\infty = 1$$

Third, if  $x \in \mathbb{Z}$ , we have

$$\begin{aligned}
 \Gamma(x + 1) &= x\Gamma(x) \\
 &= x(x - 1)\Gamma(x - 1) \\
 &\dots \\
 &= x!
 \end{aligned}$$

□

**Ex 1.18(\*\*).** *Proof.* Using the equation 5, we obtain

$$\sqrt{2\sigma^2} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} d\frac{x}{\sqrt{2\sigma^2}} = \sqrt{2\pi\sigma^2}$$

which is equal to

$$\int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi} \quad (13)$$

Otherwise, using the gamma function, we have

$$\begin{aligned} \int_0^\infty e^{-r^2} r^{D-1} dr &= \frac{1}{2} \int_0^\infty e^{-r^2} (r^2)^{D/2-1} dr^2 \\ &= \frac{1}{2} \Gamma\left(\frac{D}{2}\right) \end{aligned} \quad (14)$$

Thus, we obtain

$$\begin{aligned} S_D &= \frac{\prod_{i=1}^D \int_{-\infty}^{+\infty} e^{-x_i^2} dx_i}{\int_0^\infty e^{-r^2} r^{D-1} dr} \\ &= \frac{2\sqrt{\pi}^D}{\Gamma(D/2)} \end{aligned} \quad (15)$$

So, the surface area with radius  $r$  in  $D$  dimension wil equal to  $S_D r^{D-1}$ .

Thus, we obtain

$$V_D = \int_0^1 S_D r^{D-1} dr = \frac{S_D}{D} \quad (16)$$

□