Solutions for Chapter 5 of *Natural Image Statistics*

Yimeng Zhang

June 30, 2014

1 Mathematical Exercises

5.11.1

Let d_i denote the DC component for the ith image, $i = 1, \ldots, N$. m is number of pixels in an image.

Then

$$\begin{aligned}
d_i &= \frac{1}{m} \sum_{x',y'} I_i(x', y'), && (1) \\
E(d) &= \frac{1}{N} \sum_i \frac{1}{m} \sum_{x',y'} I_i(x', y') && (2) \\
&= \frac{1}{N} \sum_{x',y'} \frac{1}{m} \sum_i I_i(x', y') && (3) \\
&= \frac{1}{m} \sum_{x',y'} \frac{1}{N} \sum_i I_i(x', y') && (4) \\
&= \frac{1}{m} \sum_{x',y'} E_{xy} && (5) \\
&= E_{xy}, && (6) \\
E[I(x, y) - d] &= E[I(x, y)] - E(d) && (7) \\
&= E_{xy} - E_{xy} && (8) \\
&= 0, \forall x, y && (9)
\end{aligned}$$

E_{xy} is denotes $E[I(x, y)]$ that is contant for all x, y, as given by (5.52) in the book.

Note: I think chapter 5 assumes (5.52) and DC-removal for all analyses in it.

Another proof, here, I model DC component as a random variable d: mean of all pixels in a image of random pixels.

$$
\begin{aligned}
d &= \frac{1}{m}\sum_{x',y'} I_i(x',y'), && (10)\\
E(d) &= E(\sum_{x',y'} I_i(x',y')) && (11)\\
&= \frac{1}{m}\sum_{x',y'} E(I_i(x',y')) && (12)\\
&= \frac{1}{m}\sum_{x',y'} E_{xy} && (13)\\
&= E_{xy}, && (14)\\
E[I(x,y)-d] &= E[I(x,y)] - E(d) && (15)\\
&= E_{xy} - E_{xy} && (16)\\
&= 0, \forall x, y && (17)
\end{aligned}
$$

Since every feature s is just a linear combination of $\tilde{I}(x,y)$, $E(s) = 0$

5.11.2

Following (5.2) in the book, after DC removal, we have

$$
\begin{aligned}
s_i^* &= \sum_{x,y} W_i(x,y)[I(x,y)-d] && (18)\\
&= \sum_{x,y} W_i(x,y)I(x,y) - \sum_{x,y} W_i(x,y)d && (19)\\
&= s_i - 0 && (20)\\
&= s_i. && (21)
\end{aligned}
$$

s_i^* is the new feature value.

5.11.3

Trivial.

5.11.4

(a) When $E(\boldsymbol{x}) = \mathbf{0}$.
(b) Trivial. Linearity of expectation.

5.11.5

(a)

$$
\begin{aligned}
J(\boldsymbol{w}) &= E[(\boldsymbol{x} - z\boldsymbol{w})^T(\boldsymbol{x} - z\boldsymbol{w})] && (22)\\
&= E[\boldsymbol{x}^T\boldsymbol{x} - 2z\boldsymbol{w}^T\boldsymbol{x} + z^2\boldsymbol{w}^T\boldsymbol{w}] && (23)\\
&= E[\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{w}\boldsymbol{x}^T\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{w}] && (24)\\
&= \sum_j \text{var}(x_j) + E[-2\boldsymbol{x}^T\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{x} + \boldsymbol{w}^T\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{w}] && (25)\\
&= \sum_j \text{var}(x_j) + (\sum_j w_j^2)E[\boldsymbol{w}^T\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{w}] - 2E[\boldsymbol{x}^T\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{x}] && (26)\\
&= \sum_j \text{var}(x_j) + (\sum_j w_j^2)\boldsymbol{w}^T E[\boldsymbol{x}\boldsymbol{x}^T]\boldsymbol{w} - 2\boldsymbol{w}^T E[\boldsymbol{x}\boldsymbol{x}^T]\boldsymbol{w} && (27)\\
&= \sum_j \text{var}(x_j) + (\sum_j w_j^2)\sum_{i,i'} w_i w_{i'}\text{cov}(x_i, x_{i'}) - 2\sum_{i,i'} w_i w_{i'}\text{cov}(x_i, x_{i'}) && (28)
\end{aligned}
$$

$E[\boldsymbol{x}^T\boldsymbol{x}] = \sum_j \text{var}(x_j)$ because we assume mean of every component of $\boldsymbol{x}$ is zero. With $\sum_j w_j^2 = 1$, we have

$$
\begin{aligned}
J(\boldsymbol{w}) &= \sum_j \text{var}(x_j) + (\sum_j w_j^2)\sum_{i,i'} w_i w_{i'}\text{cov}(x_i, x_{i'}) - 2\sum_{i,i'} w_i w_{i'}\text{cov}(x_i, x_{i'}) && (29)\\
&= \sum_j \text{var}(x_j) - \sum_{i,i'} w_i w_{i'}\text{cov}(x_i, x_{i'}) && (30)\\
&= \sum_j \text{var}(x_j) - \boldsymbol{w}^T C\boldsymbol{w}. && (31)
\end{aligned}
$$

After this, we can go on following steps in chapter 12 of PRML (12.1.1, maximum variance formulation). Note that $\sum_j \text{var}(x_j)$ is constant in this case, and $\min J(\boldsymbol{w})$ is $\max \boldsymbol{w}^T C\boldsymbol{w}$.