EE 639: Multimedia Information System

Pattern Recognition

Homework 4 *Solution*

1. paper review: no definitive answer.

2. If Σ is a symmetric positive definite matrix, then Σ can be factored into LL^T, where L is a lower-triangular matrix with positive diagonal elements. Such a factorization is called the Cholesky decomposition.

   We can sampld $z \sim N(0,1)$ and choose $x = \mu + Lz$ to get $x \sim N(\mu, \Sigma)$, since

   $$E[x] = E[\mu + Lz] = E[\mu] + L \cdot E[z] = \mu + L \cdot 0 = \mu \tag{1}$$

   and

   $$\Sigma_x = E[Lz(Lz)^T] = E[Lzz^T L^T] = L \cdot E[zz^T] \cdot L^T = LL^T = \Sigma \tag{2}$$

3. Recall from multivariate calculus, if $x \in \Re^d$, $A \in \Re^{d \times d}$, and $f(x, A)$ is a real-value function, we have

   $$\begin{aligned} \frac{\partial}{\partial x} f &= \left(\frac{\partial f}{\partial x_1} \frac{\partial f}{\partial x_2} \cdots \frac{\partial f}{\partial x_d} \right) = \left(\frac{\partial f}{\partial x_i} \right)_i^T \\ \frac{\partial}{\partial A} f &= \left(\frac{\partial f}{\partial a_{ij}} \right)_{ij} \end{aligned}$$

   (a)

   $$\begin{aligned} \frac{\partial}{\partial x} x^T y &= \frac{\partial}{\partial x} \left(\sum_{i=1}^{d} x_i y_i \right) \\ &= (y_1 y_2 \ldots y_d)^T = y^T \end{aligned}$$

   (b)

   $$\begin{aligned} \frac{\partial}{\partial x} x^T A x &= \left[\frac{\partial}{\partial x_i} x^T \begin{pmatrix} \sum_j a_{1j} x_j \\ \vdots \\ \sum_j a_{dj} x_j \end{pmatrix} \right]_i^T \\ &= \left[\frac{\partial}{\partial x_i} \sum_k x_k \sum_j a_{kj} x_j \right]_i^T \\ &= \left[\frac{\partial}{\partial x_i} \sum_k \sum_j x_k a_{kj} x_j \right]_i^T \\ &= \left[\frac{\partial}{\partial x_i} \left(\sum_j (a_{ij} + a_{ji}) x_i x_j + \text{terms with no } x_i \right) \right]_i^T \end{aligned}$$

$$
\begin{aligned}
&= \left[\sum_j (a_{ij}+a_{ji})x_j\right]_i^T \\
&= x^T A + x^T A^T
\end{aligned}
$$

(c) Recall trace$(AB) = \sum_i \sum_j a_{ij}b_{ji}$. We first prove trace$(ABC) = $ trace$(CAB) = $ trace(BCA).

$$
\begin{aligned}
\text{trace}(ABC) &= \sum_i \sum_j a_{ij}(BC)_{ji} \\
&= \sum_i \sum_j a_{ij} \sum_k b_{jk}c_{ki} \\
&= \sum_i \sum_j \sum_k a_{ij}b_{jk}c_{ki} \\
&= \sum_j \sum_k \sum_i b_{jk}c_{ki}a_{ij} = \text{trace}(BCA) \\
&= \sum_k \sum_i \sum_j c_{ki}a_{ij}b_{jk} = \text{trace}(CAB)
\end{aligned}
$$

Since x^TAx is a real number, $x^TAx = \text{trace}(x^TAx) = \text{trace}(xx^TA)$:

$$
\begin{aligned}
\frac{\partial}{\partial A} x^TAx &= \left[\frac{\partial}{\partial a_{ij}} \text{trace}(xx^Ta)\right]_{ij} \\
&= \left[\frac{\partial}{\partial a_{ij}} \left(\sum_i \sum_j (xx^T)_{ij}a_{ji}\right)\right]_{ij} \\
&= \left[(xx^T)_{ji}\right]_{ij} \\
&= (xx^T)^T \\
&= xx^T
\end{aligned}
$$

(d)

$$
\begin{aligned}
\frac{\partial}{\partial A} \log|A| &= \left[\frac{\partial}{\partial a_{ij}} \log|A|\right]_{ij} \\
&= \frac{1}{|A|}\left[\frac{\partial}{\partial a_{ij}}|A|\right]_{ij} \\
&= \frac{1}{|A|}\left[\frac{\partial}{\partial a_{ij}} \sum_k (-1)^{i+k} a_{ik}M_{ik}\right]_{ij}
\end{aligned}
$$

where M_{ik} is the determinant of the submatrix of A with row i and column k deleted.

$$
\begin{aligned}
&= \frac{1}{|A|}\left[(-1)^{i+j}M_{ij}\right]_{ij} \\
&= A^{-1}
\end{aligned}
$$