Automatic Speech Recognition
Question Set 2

1. Given
$$Y = T\ X,$$
where X is an n-dimension random vector with mean μ_X and covariance Σ_X, T is an $m \times n$ matrix.

   (a) Show that $\mu_Y = T\mu_x$ and $\Sigma_Y = T\Sigma_X T'$, where T' is the transpose of T. Note that the covariance matrix of a random vector Z is defined as
$$\Sigma_Z = E(Z - \mu_Z)(Z - \mu_Z)',$$
where E is the expectation value operator.

   (b) Show that if X has a normal distribution, then so is Y. In other words, a linear transformation of a Gaussian vector is another Gaussian vector.

   (c) Show that if the row vectors of T consist of m eigenvectors of Σ_X with distinct eigenvalues, then Σ_Y is diagonal and the m components of Y become independent random variables.

<u>solution</u>

   (a) For the mean,
$$\mu_Y = E(TX) = T(EX) = T\mu_X.$$
For the variance,
$$\begin{aligned}\Sigma_Y &= E(Y - \mu_Y)(Y - \mu_Y)' = E[T(X - \mu_X)(T(X - \mu_X))'] \\ &= TE[(X - \mu_X)(X - \mu_X)']T' = T\Sigma_X T'.\end{aligned}$$

   (b) This is only intended to be proven in the case when $m = n$, although it is valid for $m < n$. When $m = n$, the probability density function (pdf) of Y is related to the pdf of X by the Jacobian J of the function from X to Y by
$$p_Y(y) = \frac{p_X(x)}{|J|}.$$

Here J is the determinant of the matrix whose (i,j)-element is $\frac{\partial y_i}{\partial x_j}$. For $Y = TX$,

$$\begin{aligned} p_Y(y) = \frac{p_X(x)}{||T||} &= \frac{C_X}{||T||} e^{-\frac{1}{2}(x-\mu_X)'\Sigma_X^{-1}(x-\mu_X)} \\ &= C_Y e^{-\frac{1}{2}(y-\mu_Y)'\{T\Sigma_X T'\}^{-1}(y-\mu_Y)}, \end{aligned}$$

which is indeed a normal distribution with mean μ_Y and variance $\Sigma_Y = T\Sigma_X T'$.

(c) Since the covariance matrix is symmetric, the eigenvectors corresponding to distinct eigenvalues are orthogonal. Using this property, the new covariance matrix for the Gaussian vector Y is diagonal so the components are independent. ■

2. Given N random samples, $\{X_1, X_2, \ldots, X_N\}$, the sample mean vector M is defined as

$$M = \frac{1}{N}\sum_{i=1}^{N} X_i,$$

and the (biased) sample covariance matrix is given by

$$V = \frac{1}{N}\sum_{i=1}^{N}(X_i - M)(X_i - M)',$$

These data are assumed to be independent samples of random vector X with mean μ_X and covariance Σ_X.

(a) Show that

$$E(M) = \mu_X, \text{ and } E(V) = \frac{N-1}{N}\Sigma_X.$$

(b) Suppose X is normal. Is M the maximum likelihood estimator for μ_X? How about V for Σ_X?

<u>solution</u>

(a) For the sample mean,

$$E(M) = E\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \frac{1}{N}\sum_{i=1}^{N} E(X_i) = \frac{1}{N}N\mu_X = \mu_X.$$

For the sample variance,

$$
\begin{aligned}
E(V) &= E\left(\frac{1}{N}\sum_{i=1}^{N}(X_i - M)(X_i - M)'\right) \\
&= E\left(\frac{1}{N}\sum_{i=1}^{N}[(X_i - \mu_X) - (M - \mu_X)][(X_i - \mu_X) - (M - \mu_X)]'\right) \\
&= E\left(\frac{1}{N}\sum_{i=1}^{N}(X_i - \mu_X)(X_i - \mu_X)' - (M - \mu_X)(M - \mu_X)\right) \\
&= \Sigma_X - E(M - \mu_X)(M - \mu_X)' \\
&= \Sigma_X - E\left(\frac{1}{N}\sum_{i}(X_i - \mu_X)\ \frac{1}{N}\sum_{j}(X_j - \mu_X)'\right) \\
&= \Sigma_X - \frac{1}{N}\Sigma_X = \frac{N-1}{N}\Sigma_X.
\end{aligned}
$$

(b) M is the maximum-likelihood estimator for μ_X. Suppose that the unknown mean is μ. The log data-likelihood is given by

$$L(\mu) = \log\prod_i p(X_i|\mu) = \sum_i \log p(X_i|\mu).$$

Substituting the Gaussian pdf, the terms depending on μ is

$$\sum_i -\frac{1}{2}(X_i - \mu)'P(X_i - \mu),$$

where $P = \Sigma_X^{-1}$ is the precision matrix. Setting the derivative with respect to μ to 0 will yield

$$P\sum_i (X_i - \mu) = 0,$$

and it follows that

$$\mu^* = \frac{1}{N}\sum_i X_i.$$

It can be shown that V is the ML estimate. The proof is simple in one-dimensional case but rather involved for multi-variate cases.