Statistics 324 | Winter Quarter 2008
Multivariate Analysis | Final Exam Solutions

Problem 1. Let

$$V_{n\times p} = (v_1, v_2, \ldots, v_p) \ \text{ and } \ V^-_{p\times n} = \begin{pmatrix} (v_1^-)^T \\ (v_2^-)^T \\ \vdots \\ (v_p^-)^T \end{pmatrix}.$$

(a) Define the projection $\mathbb{P}_{-j}v_j$ as the orthogonal projection of the vector v_j into the linear subspace $\mathcal{L}_{-j} \equiv \mathcal{L}(v_1, v_2, \ldots, v_{j-1}, v_{j+1}, \ldots, v_p)$.

First consider the vector v_j and its orthogonal decomposition

$$v_j = \hat{v}_j + \overset{\perp}{v}_j, \quad \text{where} \tag{1}$$

- $\hat{v}_j = \mathbb{P}_{-j}v_j$ is the projection of v_j into $\mathcal{L}_{-j}$,
- and where $\overset{\perp}{v}_j = v_j - \hat{v}_j$ is the residual vector and therefore satisfies $\hat{v}_j^T\overset{\perp}{v}_j = 0$.

Now consider an orthogonal decomposition of the vector v_j^-; that is, write,

$$v_j^- = x + y, \quad \text{where} \tag{2}$$

- $x \in \mathcal{L}_{-j}$. But $V^-V = I_p$ implies that $(v_j^-)^Tv_i = 0$ for $i \neq j$; that is, $v_j^- \perp \mathcal{L}_{-j}$. This forces $x = 0$, the zero vector.
- and $y \in \mathcal{L}(\overset{\perp}{v}_j)$. Since $\mathcal{L}(\overset{\perp}{v}_j)$ is one-dimensional, we can simply write $y = \beta\overset{\perp}{v}_j$. Therefore, Eqn. (2) is simply
$$v_j^- = \beta\overset{\perp}{v}_j \quad \text{for some } \beta \in \mathbb{R}. \tag{3}$$

By again using the fact that $(v_j^-)^Tv_i = \delta_{ji}$, we have:

$$1 \ = \ (v_j^-)^Tv_j \ = \ (\beta\overset{\perp}{v}_j)^T(\hat{v}_j + \overset{\perp}{v}_j) \ = \ (\beta\overset{\perp}{v}_j)^T\overset{\perp}{v}_j = \beta||\overset{\perp}{v}_j||^2,$$

and so from Eqn (3) we deduce $\boxed{v_j^- = (1/||\overset{\perp}{v}_j||^2)\overset{\perp}{v}_j}$, which was to be shown.

(b) Let $V_{-j} = (v_1, v_2, \ldots, v_{j-1}, v_{j+1}, \ldots, v_p)$. Then the projection operator introduced in part (a) is $\mathbb{P}_{-j} \ = \ V_{-j}(V_{-j}^TV_{-j})^{-1}V_{-j}^T$, which is idempotent (ref. page A2.2 of the notes). Then we have the simple algebraic expressions:

- $\overset{\perp}{v}_j = v_j - \hat{v}_j = (I - \mathbb{P}_{-j})v_j$;
- $||\overset{\perp}{v}_j||^2 = (\overset{\perp}{v}_j)^T\overset{\perp}{v}_j = v_j^T(I - \mathbb{P}_{-j})^T(I - \mathbb{P}_{-j})v_j = v_j^T(I - \mathbb{P}_{-j})v_j$, where the equality follows from the symmetry and idempotency of $I - \mathbb{P}_{-j}$. ♦

Problem 2. The squared volume (V^2) of the parallelpiped is

$$V^2 = \prod_{i=1}^{p} ||\overset{\perp}{u}_i||^2,$$

where $\overset{\perp}{u}_i$ is the Gram-Schmidt projection of u_i out of $\mathcal{L}(u_1, u_2, \ldots, u_{i-1})$. Because of the uniform directional distribution of each u_i, we have:

$$\mathbb{E}\left\{||\overset{\perp}{u}_i||^2\right\} = \frac{p-i+1}{p}.$$

(Note that $\overset{\perp}{u}_i$ is orthogonal to an $i-1$ dimensional subspace). The tetrahedron occupies $(1/p!)$ of the parallelpiped (Homework A3.3), so its expected squared volume is:

$$\frac{1}{(p!)^2}\prod_{i=1}^{p}\frac{p-i+1}{p} \quad = \quad \frac{1}{p!\,p^p} \quad \blacklozenge$$

Problem 3. Let $X \sim \mathcal{N}_{p\times n}(0, \not\Sigma \otimes I_n)$ have Gram-Schmidt decomposition $X = T_{p\times p}W_{p\times n}^T$ where $W^TW = I_p$ and T is lower triangular.

(a) The density of T can be computed via the change of variables formula:

$$\begin{aligned}
f^T(t) \quad &= \quad f^X(x)J(X\to T) \\
&= \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^T\not\Sigma^{-1}x_i\right\} J(X\to T) \\
&\overset{\text{(exponential is a scalar)}}{=} \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\sum_{i=1}^{n} x_i^T\not\Sigma^{-1}x_i\right)\right\} J(X\to T) \\
&= \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\operatorname{tr}\left(x_i^T\not\Sigma^{-1}x_i\right)\right\} J(X\to T) \\
&\overset{\text{(cyclic property of trace)}}{=} \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\operatorname{tr}\left(\not\Sigma^{-1}x_ix_i^T\right)\right\} J(X\to T) \\
&= \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\not\Sigma^{-1}\sum_{i=1}^{n} x_ix_i^T\right)\right\} J(X\to T) \\
&= \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\not\Sigma^{-1}XX^T\right)\right\} J(X\to T) \\
&= \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\not\Sigma^{-1}TW^TWT^T\right)\right\} J(X\to T) \\
&\overset{(W^TW=I_p)}{=} \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\not\Sigma^{-1}TT^T\right)\right\} J(X\to T) \\
&\overset{\text{(cyclic property of trace)}}{=} \quad \frac{1}{(2\pi)^{np/2}|\not\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(T^T\not\Sigma^{-1}T\right)\right\} J(X\to T).
\end{aligned}$$

But from Lemma 1 on page B2.3, we have

$$J(X \to T) = c_1 \prod_{i=1}^{p} t_{ii}^{n-i},$$

and so continuing the previous calculation, we then see,

$$f^T(t) = f^X(x) J(X \to T) = c \frac{(2\pi)^{-np/2}}{|\Sigma|^{n/2}} \prod_{i=1}^{p} t_{ii}^{n-i} e^{-\frac{1}{2}\operatorname{tr}(T^T \Sigma^{-1} T)},$$

as claimed.

(b) Now suppose $\Sigma = I_p$. So the trace term in the exponential simplifies to $\operatorname{tr}(T^T T) = \sum_{i \le j} t_{ij}^2$, and so:

$$f^T(t) \;=\; c(2\pi)^{-np/2} \prod_{i=1}^{p} t_{ii}^{n-i} e^{-\frac{1}{2}\sum_{i \le j} t_{ij}^2} \;=\; c \left(\prod_{i=1}^{p} t_{ii}^{n-i} e^{-\frac{1}{2} t_{ii}^2} \right) \left(\prod_{i<j} \right) e^{-\frac{1}{2} t_{ij}^2},$$

and so the elements of T factor, implying mutual independence of the t_{ij}'s. In particular,

- For $i < j$, the density of the off-diagonal elements t_{ij} is proportional to $e^{-\frac{1}{2} t_{ij}^2}$, indicating that they are standard Gaussian, i.e, $\mathcal{N}(0, 1)$.
- For $i = j$, the density of t_{ii}^2 is proportional to $(t_{ii}^2)^{\frac{n-i+1-1}{2}} e^{-\frac{1}{2} t_{ii}^2}$, which we recognize as $\chi^2_{(n-i+1)}$. ◆

Problem 4. Write $S_{2\times 2} = \begin{pmatrix} \sum(x_i - \bar{x})^2 & \sum(x_i - \bar{x})(y_i - \bar{y}) \\ \sum(x_i - \bar{x})(y_i - \bar{y}) & \sum(y_i - \bar{y})^2 \end{pmatrix} = LDL^T$, where L is orthonormal and D is diagonal.

(a) The first column of L (that is, the first principal component which we denote ℓ_1) and $(\bar{x}, \bar{y})^T$, describe the "least orthogonal errors" line given by:

$$\left\{ \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} + c\ell_1 \right\}, \qquad -\infty < c < \infty.$$

(b) The residual sum of orthogonal errors is d_2, the second eigenvalue of S (which is equivalent to c_2^2, where c_2 is the second singular value). ◆

Problem 5. Here, we consider the random p-vector X, which can be partitioned as

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right).$$

(a) Suppose X is normally distributed with the parameters given above. Now consider the matrix

$$A = \begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix},$$

which implies

$$AX = \begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 - \Sigma_{21}\Sigma_{11}^{-1} X_1 \end{pmatrix} \equiv \begin{pmatrix} X_1 \\ \overset{\perp}{X}_2 \end{pmatrix},$$

which has mean and variance

- $A\mu = \begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \overset{\perp}{\mu}_2 \end{pmatrix};$
- $A\Sigma A^T = \begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \overset{\perp}{\Sigma}_{22} \end{pmatrix}.$

Then recalling that the Mahalanobis distance for a Gaussian distribution is the same as its Kullback-Leibler divergence, it then follows from the *preservation property*,

$$\Delta^2(X) = \Delta^2(AX) = (\mu_1^T, \overset{\perp}{\mu}_2^{T}) \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \overset{\perp}{\Sigma}_{22}^{-1} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \overset{\perp}{\mu}_2 \end{pmatrix} = \mu_1^T \Sigma_{11}^{-1} \mu_1 + \overset{\perp}{\mu}_2^{T} \overset{\perp}{\Sigma}_{22}^{-1} \overset{\perp}{\mu}_2,$$

proving the identity.

(b) Testing $H_0 : \Delta^2 = \mu_1^T \Sigma_{11}^{-1} \mu_1$ is equivalent to testing $H_0 : \overset{\perp}{\mu}_2^{T} \overset{\perp}{\Sigma}_{22}^{-1} \overset{\perp}{\mu}_2 = 0$, or simply:

$$H_0 : \overset{\perp}{\mu}_2 = 0.$$

This hypothesis suggests using a Hotelling's T^2-type idea based on the statistic $\overset{\bar{\perp}}{X}_2$. But be careful: since we do not know any of the parameters, we have to estimate the covariance terms. Let

$$\begin{aligned} \overset{\perp}{X}_{2,i} &= X_{2,i} - S_{21} S_{11}^{-1} X_{1,i} \\ \overset{\bar{\perp}}{X}_{2,i} &= n^{-1} \sum_i \overset{\perp}{X}_{2,i} \\ \overset{\perp}{S}_{22} &= \sum_{i=1}^{n} (\overset{\perp}{X}_{2,i} - \overset{\bar{\perp}}{X}_{2,i})(\overset{\perp}{X}_{2,i} - \overset{\bar{\perp}}{X}_{2,i})^T \end{aligned}$$

So our test rejects for large values of

$$T^2 = \overset{\bar{\perp}}{X}_2^{T} \left(\frac{\overset{\perp}{S}_{22}}{n(n-1)} \right)^{-1} \overset{\bar{\perp}}{X} \sim \frac{p_2}{n - p_2} F_{p_2, n-p_2}. \ \blacklozenge$$

Problem 6. Drop the needle, and consider its orientation fixed. Then the probability that it intersects any of the lines is its projected length perpendicular to the lines. Furthermore, the needle can be thought of a convex object have length 1 in $\mathbb{R}^2$ ('volume' $= V = 2$). The needle itself is projected along lines uniformly distributed in $\mathbb{R}^2$. Then from Cauchy's projection formula,

$$V = \frac{2\bar{V}}{r_2(1,1)} \quad \Longrightarrow 2 = \frac{2 \times \Pr(\text{intersection})}{r_2(1,1)}.$$

Hence,

$$\begin{aligned} \Pr(\text{intersection}) &= r_2(1,1) \\ &= \frac{\Gamma(1)\Gamma(1)}{\Gamma(1/2)\Gamma(3/2)} \\ &= \frac{2}{\pi}. \ \blacklozenge \end{aligned}$$