BSCB 694: Theory of Multivariate Statistics

Homework 4

Solution to Question 3

In some cases, we are interested in examining the distribution of Y conditional on a set of exogenous variables X, where we assume that X and Y are jointly normally distributed.

1. **Assuming that the covariances of X and Y have full rank, derive the maximum likelihood estimates of the eigen-decomposition of the covariance matrix of Y given X.**

   We start by observing that so long as Σ_{XX} is of full rank,

   $$\Sigma_{Y|X} = \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY} \tag{1}$$

   and the eigenvalues of $\Sigma_{Y|X}$ are distinct, then there is a differentiable, 1:1 correspondence between the parametrization of $\Sigma_{Y|X}$ by its values and by the parametrization by its eigen-decomposition. Consequently, if $S_{Y|X}$ is a maximum likelihood estimate for $\Sigma_{Y|X}$, the eigen-decomposition of $S_{Y|X}$ is a maximum likelihood estimate for the eigen-decomposition of $\Sigma_{Y|X}$.

   Now, we observe from (1) that there is a 1:1 mapping between the parametrizations of

   $$\Sigma = \begin{bmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{bmatrix}$$

   given in terms of

   $$(\Sigma_{YY}, \Sigma_{YX}, \Sigma_{XX}) \text{ and } (\Sigma_{Y|X}, \Sigma_{YX}, \Sigma_{XX})$$

   so that by the same argument, the maximum likelihood estimate for $\Sigma_{Y|X}$ is

   $$S_{Y|X} = S_{YY} - S_{YX}S_{XX}^{-1}S_{XY}$$

   since we know each of the terms on the right hand side to be maximum likelihood estimates for their respective quantities.

2. **Argue that the sample properties of this decomposition are analogous to those for the unconditional PCA derived in class. (You do NOT need to repeat the proof given in class).**

   Let $\Sigma_{Y|X} = \Gamma\Lambda\Gamma^T$, $S_{Y|X} = GLG^T$. We begin by observing that $S_{Y|X}$ is distributed as a Wishart $W(\Sigma_{Y|X}, n-q)$ random variable where q is the dimension of X. From Homework 1 we know that

$$\sqrt{n}(S_{Y|X} - \Sigma_{Y|X}) \to N(0, \Omega)$$

since we can write $S_{Y|X}$ in terms of $\tilde{X}^T \tilde{X}$ for some $X \sim N(0, \Sigma_{Y|X})$.

Since the transformation

$$T : S_{Y|X} \to (G, L)$$

is 1:1 and differentiable we know that

$$\sqrt{n}[(G, L) - (\Gamma, \Lambda)] \to N\left(0, \frac{dT}{d\Sigma_{Y|X}}^T \Omega \frac{dT}{d\Sigma_{Y|X}}\right).$$

The calculations from class now represent an approximation to this distribution and can be followed directly.

3. **How would you modify your calculation of the degrees of freedom in Question 2 to conduct the same test conditional on X?**

   Since the isotropy test is a likelihood-ratio test, there is no difference in the degrees of freedom. However, note that the test statistic is different since the likelihood ratio remains proportional to

$$\mathrm{Tr}\hat{\Sigma}^{-1}S - \log|\hat{\Sigma}^{-1}S| - 1$$

   since we take

$$\hat{\Sigma} = \begin{bmatrix} S_{XX} & S_{XY} \\ S_{YX} & \hat{\Sigma}_{Y|X} + S_{YX}S_{XX}^{-1}S_{XY} \end{bmatrix}$$

   and this is *not* equivalent to replacing S with $S_{Y|X}$ and $\hat{\Sigma}$ with $\hat{\Sigma}_{Y|X}$ in the likelihood ratio statistic.