Covariance Models for Matrices

Models so far:

$$Y = M + E, \quad M \in \mathcal{M} = \begin{cases} \{M : m_{ij} = a_i + b_j\} \\ \{M : \text{rank}(M) = r\} \\ \{M : m_{ij} = \beta^T x_{ij}\} \end{cases}$$

$$E = \{e_{ij}\} \sim \text{iid } N(0, \sigma^2)$$

Is the covariance model reasonable?

Ex: $i = 1 \ldots m$ indexes subjects randomly sampled from pop
$j = 1 \ldots n$ " time points / repeated measurements, variables, etc

Random sampling structure suggests

$$E = \begin{pmatrix} e_1 - \\ \vdots \\ e_m - \end{pmatrix} \qquad e_1 \ldots e_m \overset{iid}{\sim} MVN(0, \Sigma)$$

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ & & \ddots & \\ & & & 1 \end{pmatrix} \quad \text{(exchangeable)}$$

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \\ \rho^2 & \rho & 1 & \ddots \end{pmatrix} \quad (AR)$$

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \ldots \\ \sigma_{21} & & \\ \vdots & & \end{pmatrix} \quad \text{unstructured (i.e. MVN model}$$

Ex: $i = 1 \ldots m$: random sample of "row objects"
$j = 1 \ldots n$ " "column objects"

Additive random effects model

$e_{ij} = a_i + b_j + \epsilon_{ij}$ $\quad \{a_i\} \sim_{iid} N(0, \sigma_a^2)$
$\quad \{b_j\} \sim_{iid} N(0, \sigma_b^2)$
$\quad \{\epsilon_{ij}\} \overset{iid}{\sim} N(0, \sigma_\epsilon^2)$

Two-way exchangeable covariance model:

$E[e_{ij} e_{ij}] = \sigma_a^2 + \sigma_b^2 + \sigma_\epsilon^2$ $\quad Cor(e_{ij}, e_{kl}) = 0$

$E[e_{ij} e_{ik}] = \sigma_a^2$ $\quad Cor(e_{ij}, e_{ik}) = \sigma_a^2 / \sigma_a^2 + \sigma_b^2 + \sigma_\epsilon^2$ (row correlation)

$E[e_{ij} e_{kj}] = \sigma_b^2$ $\quad Cor(e_{ij}, e_{kj}) = \sigma_b^2 / \ldots$ (column correlation)

What if we don't have random samples of objects?

Consider the mvN model: $E = \begin{pmatrix} \epsilon_1 - \\ \vdots \\ \epsilon_m - \end{pmatrix}$ $\quad \epsilon_1 \ldots \epsilon_m \sim iid\ N(0, \Sigma)$

Σ represents correlations among columns of E.

$E = \begin{pmatrix} e_1 - \\ \vdots \\ e_m - \end{pmatrix}$ $\quad e_1 \ldots e_m \overset{iid}{\sim} N(0, \Sigma)$

What is the "covariance matrix" for E?

E is $m \times n$. $\{Cov(e_{ij}, e_{kl})\}$ is an $m \times m \times n \times n$ matrix!

Vec operator:

$vec(E) = vec\left(\begin{pmatrix} e_1 & \cdots & e_n \\ | & & | \end{pmatrix}\right) = \begin{pmatrix} e_1 \\ | \\ e_2 \\ | \\ \vdots \end{pmatrix}$

Kronecker Product. $A \in \mathbb{R}^{m_1 \times n_1}$, $B \in \mathbb{R}^{m_2 \times n_2}$

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \\ & \end{pmatrix} \in \mathbb{R}^{m_1 m_2 \times n_1 n_2}$$

Important Identity $\quad \mathrm{vec}(AXB^T) = (B \otimes A)\,\mathrm{vec}(X)$

Back to MVN model

$E = \begin{pmatrix} e_1 \cdots \\ e_m - \end{pmatrix}$ $\{e_i\}$ iid $MVN(0, \Sigma)$

Let $B = \Sigma^{1/2}$, i.e., $BB^T = \Sigma$

then $e_i \overset{d}{=} B z_i$, $z_i \sim N(0, I)$ $\left(E[e_i e_i^T] = E[B z_i z_i^T B^T] = BB^T \right)$

$$E = \begin{bmatrix} B z_1 - \\ B z_2 - \\ \vdots \\ B z_m - \end{bmatrix} = \begin{bmatrix} z_1 - \\ z_2 - \\ \vdots \\ z_m - \end{bmatrix} (B^T_{n \times n}) = ZB^T \qquad Z = \{z_{ij}\} \overset{iid}{\sim} N(0,1)$$

$$\mathrm{vec}(ZB^T) = \mathrm{vec}(IZB^T) = (B \otimes I)\,\mathrm{vec}(Z) = (B \otimes I)\, z$$

$$\begin{aligned} \mathrm{Cov}(\mathrm{vec}(E)) &= \mathrm{Cov}(\mathrm{vec}(ZB^T)) \\ &= E\left[(\mathrm{vec}(ZB^T))(\mathrm{vec}(ZB^T))^T \right] \\ &= E\left[(B \otimes I)\, z z^T (B^T \otimes I) \right] \\ &= (B \otimes I)\, E[zz^T]\, (B^T \otimes I) = (B \otimes I)(B^T \otimes I) = BB^T \otimes I \end{aligned}$$

Result: $\underset{\sim}{e}_1 \ldots \underset{\sim}{e}_n \overset{iid}{\sim} MVN(0, \Sigma)$

$$\rightarrow E = \begin{pmatrix} e_1^T \\ \vdots \\ e_n^T \end{pmatrix} \overset{\mathcal{L}}{=} \begin{pmatrix} (Bz_1)^T \\ \vdots \\ (Bz_n)^T \end{pmatrix} \overset{\mathcal{L}}{=} ZB^T, \quad BB^T = \Sigma$$

Let $\underset{\sim}{e} = vec(E) \in \mathbb{R}^{M \times n} \rightarrow Cov(\underset{\sim}{e}) = \Sigma \otimes I$

Σ represents covariance between columns of E

Separable Covariance:

$Cov(vec(Z)) = I_{mn \times mn}, \quad E = ZB^T \rightarrow Cov(vec(E)) = BB^T \otimes I$

what about covariance among the rows?

Let $E = AZB^T$: $cov(vec(E)) = ?$

$$e = vec(E) = (B \otimes A)\, vec(Z) = (B \otimes A)\underset{\sim}{z}$$

$$Cov(\underset{\sim}{e}) = E(\underset{\sim}{e}\underset{\sim}{e}^T) = E\left[(B \otimes A)\underset{\sim}{z}\underset{\sim}{z}^T(B^T \otimes A^T)\right] =$$

$$= (B \otimes A)\, I\, (B^T \otimes A^T) = (BB^T \otimes AA^T)$$

$$= \Sigma_c \otimes \Sigma_r$$

Def: (Sep Cov Model): $\mathcal{S}^+_{M_1, M_2} = \{\Sigma \in \mathcal{S}^+_{M_1M_2 \times M_1M_2} : \Sigma = \Sigma_c \otimes \Sigma_r,\ \Sigma_c \in \mathcal{S}^+_{M_2},\ \Sigma_r \in \mathcal{S}^+_{M_1}\}$

Def: (Matrix normal model):

$Y \sim N_{M_1 \times M_2}(M, \Sigma_c \otimes \Sigma_r) \Leftrightarrow$ Y_{ij} are normal

$E(v(Y)) = v(M)$

Some basic properties: $Y \sim N_{n_r \times n_c}(0, \Sigma_c \otimes \Sigma_r)$

$$\Rightarrow \quad E[YY'] = \Sigma_r \, tr(\Sigma_c)$$

$$E[Y'Y] = \Sigma_c \, tr(\Sigma_r)$$

$$E[\|Y\|^2] = tr(\Sigma_r) tr(\Sigma_c)$$

$E\Lambda$ ( )(γ

$E[YY^T] = E[AZB^TBZA^T] = A\,E[ZB^TBZ]A^T.$

$= A\,(I\,tr(BB^T))\,A^T = AA^T\,tr(BB^T)$

d 1: $E[ZB^TBZ] = E[ZGZ^T] = E[Z\overset{\text{orthogonal}}{U}\Lambda U^TZ^T]$

let $\tilde{Z} = ZU \rightarrow cov(vec(\tilde{Z})) = E[U^T zz^T U] = U'U = I$

$= E[\tilde{Z}\Lambda\tilde{Z}^T]$

z_i = ith row of Z

$= E\begin{bmatrix} z_1^T\Lambda z_1 & z_1^T\Lambda z_2 & \cdots \\ \vdots & & \\ \vdots & & \end{bmatrix}$

$E[z_i^T\Lambda z_i] = E[\sum \lambda_j z_j^2] = tr(\Lambda)$

$E[z_i^T\Lambda z_j] = 0$ if $i \neq j$

$tr(\Lambda)\,I$

but $tr(\Lambda) = tr(U^TU\Lambda)$

$= tr(U\Lambda U^T)$

$= tr(G)$

$= tr(B^TB)$