Indexing

Indexing Units:

N_0 = # training households
N_1 = # testing households (to be ranked)
K = # communities

Indexing Information:

A = # of response variables collected at community-level
M = # of response variables collected at household-level
R = # of rankers of the N_1 test units/households
P = # of covariates

Inputs:

$$y^{comm} = \left[y_1^{comm}, \dots, y_A^{comm}\right]_{K \times A}$$ community-level

$$y^{micro} = \left[y_1^{micro}, \dots, y_M^{micro}\right]_{N_0 \times M}$$ training units

$$T = \left[T(\underset{\sim}{z}_1), \dots, T(\underset{\sim}{z}_R)\right]_{N_1 \times R}$$ testing units

X_0^{micro} is a $N_0 \times P$ matrix of covariates training units
X_1^{micro} is a $N_1 \times P$ matrix of covariates testing units

X^{comm} is a $K \times P$ matrix of covariates community level

→ Q: X^{comm} is just a mean-aggregation of $\begin{bmatrix} X_0 \\ X_1 \end{bmatrix}$, right?

Re-structuring of Inputs:

Take the vec( ) function on a $R \times C$ matrix
$X = [\underset{\sim}{x}_1, \dots, x_C]$ to be column-wise stacking.
$vec(X) = (\underset{\sim}{x}_1^T, \underset{\sim}{x}_2^T, \dots, \underset{\sim}{x}_C^T)^T \leftarrow (RC) \times 1$ vector

$$u = \begin{pmatrix} vec(y^{comm}) \\ vec(y^{micro}) \\ vec(Z) \end{pmatrix}$$ is a $(AK + MN_0 + RN_1) \times 1$ vector

$$X = \begin{bmatrix} I_{A \times 1} \otimes X^{comm} \\ I_{M \times 1} \otimes X_0^{micro} \\ I_{R \times 1} \otimes X_1^{micro} \end{bmatrix}$$ is a $(AK + MN_0 + RN_1) \times P$ matrix

$$\begin{pmatrix} X^{comm} & & & & 0 \\ & X^{com} & & & \\ & & \ddots & X_j^{micro} & \\ 0 & & & \ddots & X'^{micro} \end{pmatrix}$$

Model

This is a 1st stage model - we will have to build upon it later (e.g. random effects, etc...)

community random effects not captured by $X\beta$

For $a = 1, \dots, A$; $k = 1, \dots, K$: $y_{a,k}^{comm} \sim N\left(\gamma_k^{comm} + \alpha_a^{comm} + (\underset{\sim}{X}_k^{comm})^T \underset{\sim}{\beta}, (w_a^{comm})^{-1}\right)$

For $m = 1, \dots, M$; $i = 1, \dots, N_0$: $y_{m,i}^{micro} \sim N\left(\gamma_i^{micro} + \alpha_m^{micro} + (X_{0i}^{micro})^T \underset{\sim}{\beta}, (w_k^{micro})^{-1}\right)$

For $r = 1, \dots, R$; $j = 1, \dots, N_1$: $z_{rj} \sim \left(\gamma_j^{z} + \alpha_r + (X_{1j}^{micro})^T \underset{\sim}{\beta}, (w_r^{rank})^{-1}\right)$ testing latent scores

what's left over for j-th test subject after $X_j\beta$, expecting variability $\frac{1}{w_r^{rank}}$. (e.g. rankers tended to rank unit j higher than $X\beta$ suggested)

→ equivalently,

$$u \sim MVN(X^R \underset{\sim}{\gamma} + \underset{\sim}{\alpha} \underset{\sim}{1} + X^F \underset{\sim}{\beta}, \Sigma)$$

where

$$\Sigma^{-1} = diag\left((\underset{\sim}{w}^{com} \otimes \underset{\sim}{1})^T, (\underset{\sim}{w}^{micro} \otimes \underset{\sim}{1}), (\underset{\sim}{w}^{rank} \otimes \underset{\sim}{1})\right)$$

Priors:

$\alpha \sim N(0, \sigma_\alpha^2)$ fix at first
$\underset{\sim}{\beta} \sim MVN(0, \sigma_\beta^2 I_{P \times P})$

$$w_a^{comm} = \begin{cases} 0.5 & w.p. \ 1/3 \\ 1 & w.p. \ 1/3 \\ 2 & w.p. \ 1/3 \end{cases}$$

$$w_m^{micro} = \begin{cases} 0.5 & w.p. \ 1/3 \\ 1 & w.p. \ 1/3 \\ 2 & w.p. \ 1/3 \end{cases}$$

$$w_r^{rank} = \begin{cases} 0.5 & w.p. \ 1/3 \\ 1 & w.p. \ 1/3 \\ 2 & w.p. \ 1/3 \end{cases}$$

larger w ⇕ that measure has "stronger" relationship w/ X covariates. not quite right.

smaller w ⇕ that measure isn't described by $\alpha + X\beta$ as the other measures are. → less reliable?

Random effects for entities

Variability attributed to entities/households not explained by X^F

$$\underset{\sim}{u} = X^R \underset{\sim}{\gamma} + X^F \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

↑ design matrix for random effects

↑ design matrix for fixed effects

Revised Posteriors - β

$$f(\underset{\sim}{\beta} | \cdot) \propto \exp\left\{-\frac{1}{2}\left[-2(\underset{\sim}{u} - X^R \underset{\sim}{\gamma})^T \Sigma^{-1} X^F \underset{\sim}{\beta} + \underset{\sim}{\beta}^T \left(X^{FT} \Sigma^{-1} X^F + \frac{1}{\sigma_\beta^2} I\right) \underset{\sim}{\beta}\right]\right\}$$

$$\rightarrow \beta | \cdot \sim MVN_{P+1}(\bar{\mu}_\beta, \bar{V}_\beta)$$

$$\bar{V}_\beta = \left(X^{FT} \Sigma^{-1} X^F + \frac{1}{\sigma_\beta^2} I\right)^{-1}$$

$$\bar{\mu}_\beta = (\underset{\sim}{u} - X^R \underset{\sim}{\gamma})^T \Sigma^{-1} X^F \left(X^{FT} \Sigma^{-1} X^F + \frac{1}{\sigma_\beta^2} I\right)^{-1}$$

Posterior for $\underset{\sim}{\gamma}$ (: do separate for micro, comm, Z)

$$\gamma | \cdot \sim MVN(\bar{\mu}_\gamma, \bar{V}_\gamma)$$

$$\bar{V}_\gamma = \left(X^{RT} \Sigma^{-1} X^R + \frac{1}{\sigma_\gamma^2} I\right)^{-1}$$

$$\bar{\mu}_\gamma = \boxed{\underbrace{(\underset{\sim}{u} - X^F \underset{\sim}{\beta})^T}_{RN_1 \times 1 \quad RN \times 1} \Sigma^{-1} X^R}\left(X^{RT} \Sigma^{-1} X^R + \frac{1}{\sigma_\gamma^2} I\right)^{-1}$$

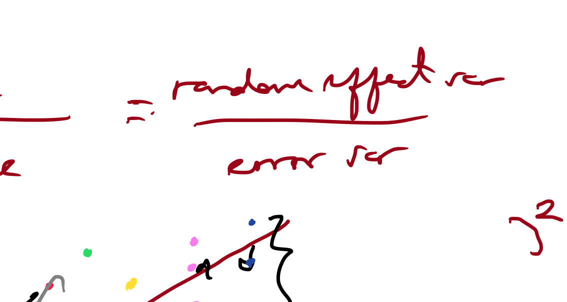
$1 \times RN \quad RN \times RN$

$1 \times 2N$

Q: what if $\frac{\sigma_{rank}^2}{(1/w_{rank})} = \frac{\text{random effect var}}{\text{error var}}$ is large

variance of $\{\gamma_j\}_{j=1}^{N_1}$

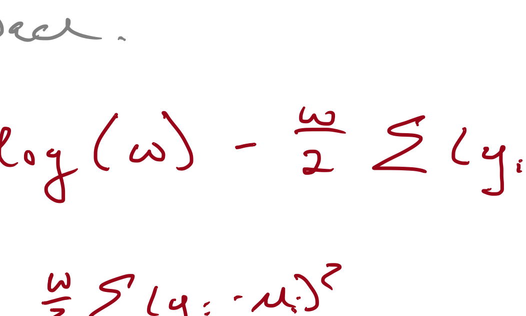
ranks of all "test" subjects from R rankers → Z



→ Ranker consistency and/or, some ranker disagreement with the space that can be explained by X^F

Q: what if $\frac{\sigma_{rank}^2}{1/w_{rank}} = \frac{\text{random effect var}}{\text{error var}}$ is small?

w/in same person, lots of rank variability



γ^2

→ ranker-to-ranker disagreement (error big) random $\frac{1}{w_{rank}}$ and/or rankers, on avg, tend to agree w/ X space.

$$\log(\text{prior}) + \frac{n}{2}\log(\omega) - \frac{\omega}{2}\sum(y_i - \mu$$

$$\text{prior} \times \omega^{n/2} \times e^{\frac{\omega}{2}\sum(y_i - \mu_i)^2}$$

$$\text{prior} \times \prod_{i=1}^{n}\left(\sqrt{\omega}\, e^{\frac{\omega}{2}(y_i - \mu_i)^2}\right)$$

$$\text{prior} \times \prod_{i=1}^{n}\left(\frac{1}{\sqrt{\omega^{-1}}} e^{\frac{1}{2 \cdot 1/\omega}(y_i - \mu_i)^2}\right)$$