Finite mixtures of matrix normal distributions for classifying three-way data.

For one time point:

$Y = [\quad] \to n$ observations

$\downarrow$

r covariates.

or For one observation

$Y = [\quad] \longrightarrow p$ time point

$\downarrow$

r covariates

Matrix normal distribution:

$$\underset{r \times p}{Y} \sim MND(\underset{r\times p}{M}, \overset{\text{column covariance}}{\underset{r\times r}{\Phi}}, \overset{\text{row covariance}}{\underset{p\times p}{\Omega}})$$

$$\phi_i = f(Y \mid M, \Phi, \Omega) = (2\pi)^{-\frac{rp}{2}} |\Phi|^{-\frac{p}{2}} |\Omega|^{-\frac{r}{2}} \exp\left\{\frac{1}{2} tr\left[\Phi^{-1}(Y-M)\,\Omega^{-1}(Y-M)^T\right]\right\}$$

$\downarrow$ to multivariate normal:

$\Sigma = \Phi \otimes \Omega$. $\mu = vec(M)$. $vec(Y) \sim MVN(\mu, \Sigma)$

Assume our population has k sub-population, give each sub a prior π_i $\quad \sum \pi_i = 1$

Likelihood: $\sum_{i=1}^{k} \pi_i \phi_i(Y; M_i, \Phi_i, \Omega_i)$.

Set $z_{ij} = I(Y_{ij} \in i\text{th component})$

π_{ij} is posterior ~~por~~ probability: $P((i,j)$ falls into ith component $\mid Y = y)$