Greedy PMF.

We want to solve this model iteratively:

$$\boxed{X_{ij} \sim Pois(\lambda_{ij} + l_{ik} f_{jk})}$$

The MLE problem:

$$\boxed{\begin{aligned} \max_{l_{ik}, f_{jk} \geq 0} J^{(k)} &:= \sum_{ij} -(\lambda_{ij} + l_{ik} f_{jk}) + X_{ij} \log(\lambda_{ij} + l_{ik} f_{jk}) + const \\ &= \underbrace{\sum_{ij} -l_{ik} f_{jk} + X_{ij} \log(\lambda_{ij} + l_{ik} f_{jk})}_{(*)} + const \end{aligned}}$$

Note: $(*)$ is convex w.r.t l_{ik} when f_{jk} is fixed, but doesn't have analytic solution for $l_{ik}^*(f_{jk})$.

By Jensen Inequality:

$$(*) \geq \sum_{ij} -l_{ik} f_{jk} + X_{ij} \left[(1 - \xi_{ijk}) \log\left(\frac{\lambda_{ij}}{1 - \xi_{ijk}}\right) + \xi_{ijk} \log\left(\frac{l_{ik} f_{jk}}{\xi_{ijk}}\right) \right]$$

and "=" iff $\xi_{ijk} = \dfrac{l_{ik} f_{jk}}{l_{ik} f_{jk} + \lambda_{ij}}$

Note the RHS is just $\bar{Z}_{ijk} \sim Pois(l_{ik} f_{jk})$

where $\bar{Z}_{ijk} = X_{ij} \cdot \xi_{ijk}$.

So we know the optimality condition for $(*)$:

$$
\begin{cases}
\xi_{ijk} = \dfrac{l_{ik} f_{jk}}{\lambda_{ij} + l_{ik} f_{jk}} \\
l_{ik} = \dfrac{\sum_j x_{ij} \xi_{ijk}}{\sum_j f_{jk}} \\
f_{jk} = \dfrac{\sum_i x_{ij} \xi_{ijk}}{\sum_i l_{ik}}
\end{cases}
$$

In fact we use the opt-condition as coordinate updates as they also satisfy opt for each coordinate in (*)

Algo: Greedy PMF (X, K)

Init: λ_{ij}, L, F (from rank-1 PMF ; $col(L) = col(F) = 1$)

For $k \in [K]$:

$(l_{ik}, f_{jk}) \leftarrow$ rank1-update(X_{ij}, λ_{ij})

$\lambda_{ij} \leftarrow \lambda_{ij} + l_{ik} f_{jk}$

$L \leftarrow$ cbind(L, l_{Ik})

$F \leftarrow$ cbind(F, f_{Jk})

If test_terminate(l_{Ik}, f_{Jk}) is TRUE :

Break

Return: L, F.

Algo: rank1-update (X, Λ) :

Init: $(l_i, f_j) \leftarrow$ init(X, Λ).

while not converged :

$$\xi_{ij} \leftarrow \frac{l_i f_j}{l_i f_j + \lambda_{ij}}$$

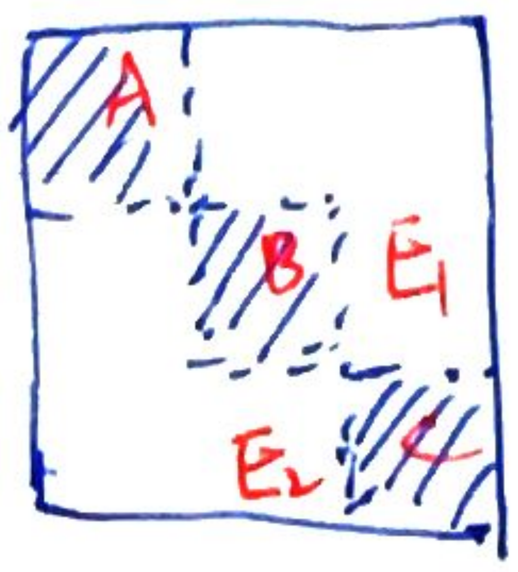
$$l_i \leftarrow \frac{\sum_j X_{ij} \xi_{ij}}{\sum_j f_j}$$

$$f_j \leftarrow \frac{\sum_i X_{ij} \xi_{ijk}}{\sum_i l_i}$$

Return (l_i, f_j)

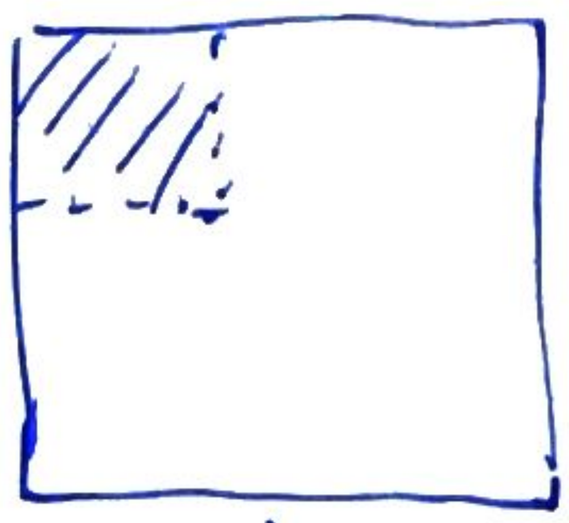
Problem with greedy PMF.

Example: block-structure.



X_{ij}

If we are lucky that λ_{ij} captures structure in A:

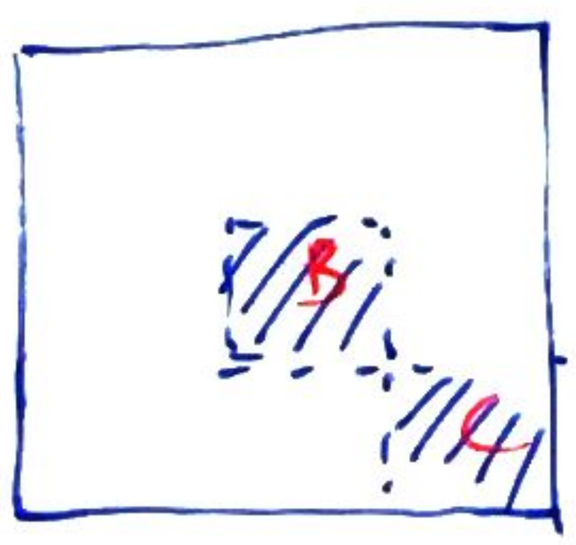


λ_{ij}

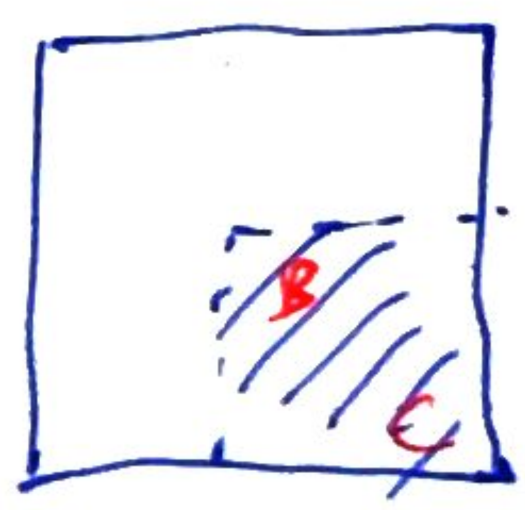
then $\lambda_{ij} \approx 0$ in $\neg A$.

this means $\xi_{ijk} = \dfrac{l_{ik} f_{jk}}{\lambda_{ij} + l_{ik} f_{jk}} \approx 1$ in $\neg A$

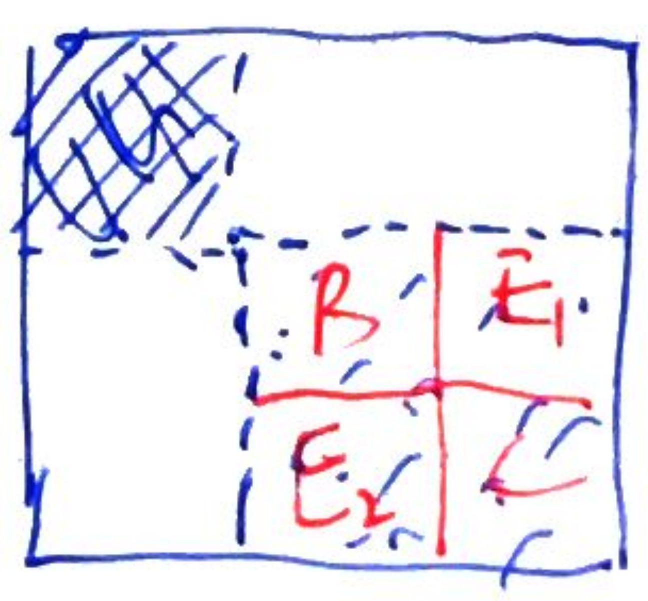
Then l_{ik}, f_{jk} would be close to col/row sum (proportional) of



will result in $l_{ik} f_{jk}$ being



after updating λ_{ij}, we have



λ_{ij}.

There is no way we can fix the error in E_1, E_2.