

2)

$$\text{tr} \left[ D_r^{-\frac{1}{2}} (P - \hat{P}) D_c^{-1} (P - \hat{P})^t D_r^{-\frac{1}{2}} \right] = \text{tr} \left[ \left( D_r^{-\frac{1}{2}} (P - \hat{P}) D_c^{-\frac{1}{2}} \right) \left( D_r^{-\frac{1}{2}} (P - \hat{P}) D_c^{-\frac{1}{2}} \right)^t \right]$$

By SVD,  $D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}} = U \Lambda V^t$ .

By Eckart-Young Theorem, the best t-rank reduced matrix  $D_r^{-\frac{1}{2}} \hat{P} D_c^{-\frac{1}{2}} = U \Lambda_t V^t$ , where the first t diagonal elements of  $\Lambda_t$  are same with those of  $\Lambda$  and other elements are zeros. Thus,  $\hat{P} = D_r^{\frac{1}{2}} U \Lambda_t V^t D_c^{\frac{1}{2}}$ .

When  $t = 1$ ,  $D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}} = \left( \frac{n_{ij}}{\sqrt{n_i} \sqrt{n_j}} \right)_{ij}$ . Meanwhile we have  $D_r^{-\frac{1}{2}} r = \left( \sqrt{\frac{n_{i+}}{n}} \right)_i$  and  $D_c^{-\frac{1}{2}} c = \left( \sqrt{\frac{n_{+j}}{n}} \right)_j$ .

$$\begin{aligned} \left( D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}} \right) \left( D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}} \right)^t \left( D_r^{-\frac{1}{2}} r \right) &= \left( \sum_{k,j} \frac{n_{ik}}{\sqrt{n_{i+} n_{+k}}} \frac{n_{jk}}{\sqrt{n_{j+} n_{+k}}} \sqrt{\frac{n_{j+}}{n}} \right)_i = \left( \sqrt{\frac{n_{i+}}{n}} \right)_i = D_r^{-\frac{1}{2}} r \\ \left( D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}} \right)^t \left( D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}} \right) \left( D_c^{-\frac{1}{2}} c \right) &= \left( \sum_{k,i} \frac{n_{kj}}{\sqrt{n_{k+} n_{+j}}} \frac{n_{ki}}{\sqrt{n_{k+} n_{+i}}} \sqrt{\frac{n_{+i}}{n}} \right)_j = \left( \sqrt{\frac{n_{+j}}{n}} \right)_j = D_c^{-\frac{1}{2}} c \end{aligned}$$

So we know  $D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}}$  have singular value  $\lambda = 1$ . We denote  $D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}}$  as  $A$  and we know its element  $\frac{n_{ij}}{\sqrt{n_{i+} n_{+j}}} \leq 1$ . So  $\lim_n (A A^t)^n = B$  and each row and column of B is either 1 or 0. If A had a singular value greater than 1, this would not happen. So  $\lambda = 1$  is the greatest singular value. Namely,

$$D_r^{-\frac{1}{2}} P D_c^{-\frac{1}{2}} = (D_r^{-\frac{1}{2}} r | U_{n*(r-1)}) \begin{pmatrix} 1 & 0 \\ 0 & \Lambda' \end{pmatrix} (D_c^{-\frac{1}{2}} c | V_{n*(s-1)})^t$$

When rank  $t = 1$ ,  $D_r^{-\frac{1}{2}} \hat{P} D_c^{-\frac{1}{2}} = D_r^{-\frac{1}{2}} r c^t D_c^{-\frac{1}{2}} \Rightarrow \hat{P} = r c^t$ .