# *Discrete Personalized Ranking* 阅读报告

这一篇论文提出的模型与 *DCF* 基本一致， 不同的地方在于使用了 AUC 作为目标函数进行优化。因此，本篇报告主要是对**四个子问题**以及**初始化**问题的推导。

- **objective function**

$$
\begin{gathered}
\underset{B,D,X,Y}{\arg\min} \sum_{(u,i,j)\in D_S} \frac{1}{|U||I_u^+||I_u^-|}\left(2r - \mathbf{b}_u^T(\mathbf{d}_i - \mathbf{d}_j)\right)^2 \\
- 2\alpha\, tr(\mathbf{B}^T\mathbf{X}) - 2\beta\, tr(\mathbf{D}^T\mathbf{Y}) \\
s.t.\ \mathbf{B} \in \{\pm 1\}^{r\times n}, \mathbf{D} \in \{\pm 1\}^{r\times m} \\
\mathbf{X}\mathbf{1}_n = 0,\ \mathbf{Y}\mathbf{1}_m = 0,\ \mathbf{X}\mathbf{X}^T = n\mathbf{I}_r,\ \mathbf{Y}\mathbf{Y}^T = m\mathbf{I}_r
\end{gathered}
$$

- **B-subproblem**

$$
\begin{aligned}
\underset{b_{uk}\in\pm 1}{\arg\min} \sum_{i,j\in I} z_u^+ z_u^- r_{ui}(1 - r_{uj})\Big(&((\mathbf{d}_i - \mathbf{d}_j)^T\mathbf{b}_u)^2 \\
&- 4r(\mathbf{d}_i - \mathbf{d}_j)^T\mathbf{b}_u\Big) - 2\alpha n \mathbf{x}_u^T\mathbf{b}_u
\end{aligned} \tag{1}
$$

其中

$$
\begin{aligned}
((\mathbf{d}_i - \mathbf{d}_j)^T\mathbf{b}_u)^2 = &\underbrace{((\mathbf{d}_{i\bar{k}} - \mathbf{d}_{j\bar{k}})^T\mathbf{b}_{u\bar{k}})^2 + ((d_{ik} - d_{jk})b_{uk})^2}_{\text{constant}} \\
&+ 2(\mathbf{d}_{i\bar{k}} - \mathbf{d}_{j\bar{k}})^T\mathbf{b}_{u\bar{k}}(d_{ik} - d_{jk})b_{uk}
\end{aligned}
$$

$$
(\mathbf{d}_i - \mathbf{d}_j)^T\mathbf{b}_u = \underbrace{(\mathbf{d}_{i\bar{k}} - \mathbf{d}_{j\bar{k}})^T\mathbf{b}_{u\bar{k}}}_{constant} + (d_{ik} - d_{jk})b_{uk}
$$

$$
\mathbf{x}_u^T\mathbf{b}_u = \underbrace{\mathbf{x}_{u\bar{k}}^T\mathbf{b}_{u\bar{k}}}_{constant} + x_{uk}b_{uk}
$$

故式 (1) 等价于：

$$
\begin{aligned}
\underset{b_{uk}\in\pm 1}{\arg\min} \sum_{i,j\in I} z_u^+ z_u^- r_{ui}(1 - r_{uj})\Big(&(\mathbf{d}_{i\bar{k}} - \mathbf{d}_{j\bar{k}})^T\mathbf{b}_{u\bar{k}}(d_{ik} - d_{jk})b_{uk} \\
&-2r(d_{ik} - d_{jk})b_{uk}\Big) - \alpha n x_{uk}b_{uk}
\end{aligned} \tag{2}
$$

令

$$
\begin{aligned}
\hat{b}_{uk} = \sum_{i,j\in I} z_u^+ z_u^- r_{ui}(1 - r_{uj})\Big(&(\mathbf{d}_{i\bar{k}} - \mathbf{d}_{j\bar{k}})^T\mathbf{b}_{u\bar{k}}(d_{ik} - d_{jk}) \\
&- 2r(d_{ik} - d_{jk})\Big) - \alpha n x_{uk}
\end{aligned}
$$

式 (2) 可以写成

$$\arg\min_{b_{uk}\in\pm1} b_{uk}\hat{b}_{uk}$$

- **D-subproblem**

  与 **B-subproblem** 类似，不过似乎不能像 B-subproblem 一样并行计算。

- **X-subproblem**

$$\arg\max_{\mathbf{X}\in\mathbb{R}^{r\times n}} tr(\mathbf{B}^T\mathbf{X}),\ s.t.\ \mathbf{X1}=0,\ \mathbf{XX}^T=n\mathbf{I} \tag{3}$$

  论文给出了该问题的解为：

$$\mathbf{X}^*=\sqrt{n}[\mathbf{P}_b\ \hat{\mathbf{P}}_b][\mathbf{Q}_b\ \hat{\mathbf{Q}}_b]^T \tag{4}$$

  - $[\mathbf{P}_b\ \hat{\mathbf{P}}_b]$ 可以通过对 $\overline{\mathbf{B}}\overline{\mathbf{B}}^T$ 进行特征值分解得到

$$\overline{\mathbf{B}}\overline{\mathbf{B}}^T=[\mathbf{P}_b\ \hat{\mathbf{P}}_b]\begin{bmatrix}\mathbf{\Sigma}_b^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}[\mathbf{P}_b\ \hat{\mathbf{P}}_b]^T$$

    其中：

$$\overline{\mathbf{B}}=\mathbf{BJ},\ \mathbf{J}=\mathbf{I}-\frac{1}{n}\mathbf{11}^T$$

  - $\mathbf{Q}_b\in\mathbb{R}^{m\times r'}$ 可以通过 SVD 的定义得到

$$\overline{\mathbf{B}}=\mathbf{P}_b\mathbf{\Sigma}_b\mathbf{Q}_b^T=\sum_{k=1}^{r'}\sigma_k\mathbf{p}_k\mathbf{q}_k^T\text{，其中}r'<r\text{是 }\overline{\mathbf{B}}\text{的秩，}\sigma_k\text{为奇异值}$$

$$\mathbf{Q}_b=\overline{\mathbf{B}}^T\mathbf{P}_b\mathbf{\Sigma}_b^{-1}$$

  - $\hat{\mathbf{Q}}_b$ 可以通过对 $[\mathbf{Q}_b\ \mathbf{1}]$ 进行 Gram-Schmidt 正交化得到

  下面证明 (4) 是 (3) 的最优解：

  - 先证明 (4) 是可行解，即 $\mathbf{X}^*\in\mathcal{B}=\{\mathbf{X}\in\mathbb{R}^{r\times n}|\mathbf{X1}=\mathbf{0},\mathbf{XX}^T=n\mathbf{I}\}$.

    注意到 $\mathbf{J1}=\mathbf{0}$, 因此 $\mathbf{BJ1}=\mathbf{0}$。因为 $\mathbf{BJ}$ 与 $\mathbf{Q}_b^T$ 有相同的行空间，所以$\mathbf{Q}_b^T\mathbf{1}=\mathbf{0}$ 。并且，$\hat{\mathbf{Q}}_b$ 是通过对 $[\mathbf{Q}_b\ \mathbf{1}]$ 进行 Gram-Schmidt 正交化得到的，因此$\hat{\mathbf{Q}}_b^T\mathbf{1}=\mathbf{0}$。所以我们得到了$[\mathbf{Q}_b\ \hat{\mathbf{Q}}_b]^T\mathbf{1}=\mathbf{0}$, 即 $\mathbf{X}^*\mathbf{1}=\mathbf{0}$。

    另一方面，$\mathbf{X}^*\mathbf{X}^{*T}=n[\mathbf{P}_b\ \hat{\mathbf{P}}_b][\mathbf{Q}_b\ \hat{\mathbf{Q}}_b]^T[\mathbf{Q}_b\ \hat{\mathbf{Q}}_b][\mathbf{P}_b\ \hat{\mathbf{P}}_b]^T=n\mathbf{I}_r$

  - 下面证明 (4) 是 (3) 的最优解：

    考虑任意 $\mathbf{X}\in\mathcal{B}$，由于$\mathbf{X1}=\mathbf{0}$，所以$\mathbf{XJ}=\mathbf{XI}-\frac{1}{n}\mathbf{X11}^T=\mathbf{X}$，且 $\langle\mathbf{B},\mathbf{X}\rangle=\langle\mathbf{B},\mathbf{XJ}\rangle=\langle\mathbf{BJ},\mathbf{X}\rangle$。

$$
\begin{aligned}
tr(\mathbf{B}^T\mathbf{X}^*) &= \langle \mathbf{B}, \mathbf{X}^* \rangle = \langle \mathbf{BJ}, \mathbf{X}^* \rangle \\
&= \left\langle [\mathbf{P}_b\ \hat{\mathbf{P}}_b] \begin{bmatrix} \mathbf{\Sigma}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{Q}_b\ \hat{\mathbf{Q}}_b]^T, \mathbf{X}^* \right\rangle \\
&= \sqrt{n} \left\langle \begin{bmatrix} \mathbf{\Sigma}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \mathbf{I}_n \right\rangle \\
&= \sqrt{n} \sum_{k=1}^{r'} \sigma_k
\end{aligned}
$$

根据 *Neumann's trace inequality* 以及 $\mathbf{X}\mathbf{X}^T = n\mathbf{I}_r$有，

$$
\langle \mathbf{BJ}, \mathbf{X} \rangle \le \sqrt{n} \sum_{k=1}^{r'} \sigma_k
$$

因此，对于任意$\mathbf{X} \in \mathcal{B}$:

$$
\begin{aligned}
tr(\mathbf{B}^T\mathbf{X}) &= \langle \mathbf{B}, \mathbf{X} \rangle = \langle \mathbf{BJ}, \mathbf{X} \rangle \\
&\le \sqrt{n} \sum_{k=1}^{r'} \sigma_k \\
&= tr(\mathbf{B}^T\mathbf{X}^*)
\end{aligned}
$$

- **Y-subproblem**

  与 **X-subproblem** 类似

- **初始化**
  - 目标函数

$$
\begin{aligned}
\arg\min_{\mathbf{P},\mathbf{Q},\mathbf{X},\mathbf{Y}} & \sum_{u,i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(1-\mathbf{p}_u^T(\mathbf{q}_i - \mathbf{q}_j))^2 + \\
& \alpha_1 n ||\mathbf{P}||_F^2 + \beta_1 n ||\mathbf{Q}||_F^2 - 2\alpha_2 n tr(\mathbf{P}^T\mathbf{X}) - 2\beta_2 n tr(\mathbf{Q}^T\mathbf{Y}) \\
& s.t. \mathbf{X}\mathbf{1}_n = 0, \mathbf{Y}\mathbf{1}_m = 0, \mathbf{X}\mathbf{X}^T = n\mathbf{I}_r, \mathbf{Y}\mathbf{Y}^T = m\mathbf{I}_r
\end{aligned}
\tag{5}
$$

  - $\mathbf{X}$ 和 $\mathbf{Y}$ 可以通过 **X/Y-subproblem** 的解进行更新
  - $\mathbf{P}$ 和 $\mathbf{Q}$ 基于令相应的偏导数为 0 更新，但是论文并没有给出公式。
  - 固定$\mathbf{Q}$、$\mathbf{X}$、$\mathbf{Y}$，更新$\mathbf{P}$

    考虑 $\mathbf{p}_u$：

$$
\begin{aligned}
&\arg\min_{\mathbf{p}_u} \sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(1-\mathbf{p}_u^T(\mathbf{q}_i - \mathbf{q}_j))^2 + \alpha_1 n ||\mathbf{p}_u||^2 - 2\alpha_2 n \mathbf{p}_u^T \mathbf{x}_u \\
= &\arg\min_{\mathbf{p}_u} \sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})((\mathbf{p}_u^T(\mathbf{q}_i - \mathbf{q}_j))^2 - 2\mathbf{p}_u^T(\mathbf{q}_i - \mathbf{q}_j)) + \\
&\alpha_1 n ||\mathbf{p}_u||^2 - 2\alpha_2 n \mathbf{p}_u^T \mathbf{x}_u
\end{aligned}
$$

    令

    $\mathcal{J} = \sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})((\mathbf{p}_u^T(\mathbf{q}_i - \mathbf{q}_j))^2 - 2\mathbf{p}_u^T(\mathbf{q}_i - \mathbf{q}_j)) + \alpha_1 n ||\mathbf{p}_u||^2 - 2\alpha_2 n \mathbf{p}_u^T \mathbf{x}_u$

$$
\frac{\partial \mathcal{J}}{\partial \mathbf{p}_u} = \sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(2\mathbf{p}_u^T(\mathbf{q}_i - \mathbf{q}_j)(\mathbf{q}_i - \mathbf{q}_j) - 2(\mathbf{q}_i - \mathbf{q}_j)) + 2\alpha_1 n \mathbf{p}_u - 2\alpha_2 n \mathbf{x}_u = 0
$$

化简得：

$$\sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})\mathbf{p}_u^T(\mathbf{q}_i-\mathbf{q}_j)(\mathbf{q}_i-\mathbf{q}_j)+\alpha_1 n\mathbf{p}_u=\sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(\mathbf{q}_i-\mathbf{q}_j)+\alpha_2 n\mathbf{x}_u$$
$$\sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(\mathbf{q}_i-\mathbf{q}_j)^T(\mathbf{q}_i-\mathbf{q}_j)\mathbf{p}_u+\alpha_1 n\mathbf{I}_r\mathbf{p}_u=\sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(\mathbf{q}_i-\mathbf{q}_j)+\alpha_2 n\mathbf{x}_u$$
$$\mathbf{p}_u^*=\left(\sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(\mathbf{q}_i-\mathbf{q}_j)^T(\mathbf{q}_i-\mathbf{q}_j)+\alpha_1 n\mathbf{I}_r\right)^{-1}\left(\sum_{i,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(\mathbf{q}_i-\mathbf{q}_j)+\alpha_2 n\mathbf{x}_u\right)$$

- 固定**P**、**X**、**Y**，更新**Q**

  考虑 $\mathbf{q}_i$：

$$\begin{aligned}
&\arg\min_{\mathbf{q}_i}\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(1-\mathbf{p}_u^T(\mathbf{q}_i-\mathbf{q}_j))^2\\
&+\sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})(1-\mathbf{p}_u^T(\mathbf{q}_j-\mathbf{q}_i))^2\\
&+\beta_1 n||\mathbf{q}_i||^2-2\beta_2 n\mathbf{q}_i^T\mathbf{y}_i\\
=&\arg\min_{\mathbf{q}_i}\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})((\mathbf{p}_u^T(\mathbf{q}_i-\mathbf{q}_j))^2-2\mathbf{p}_u^T(\mathbf{q}_i-\mathbf{q}_j))\\
&+\sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})((\mathbf{p}_u^T(\mathbf{q}_j-\mathbf{q}_i))^2-2\mathbf{p}_u^T(\mathbf{q}_j-\mathbf{q}_i))\\
&+\beta_1 n||\mathbf{q}_i||^2-2\beta_2 n\mathbf{q}_i^T\mathbf{y}_i
\end{aligned}$$

  令

$$\begin{aligned}
\mathcal{L}=&\arg\min_{\mathbf{q}_i}\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})((\mathbf{p}_u^T(\mathbf{q}_i-\mathbf{q}_j))^2-2\mathbf{p}_u^T(\mathbf{q}_i-\mathbf{q}_j))\\
&+\sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})((\mathbf{p}_u^T(\mathbf{q}_j-\mathbf{q}_i))^2-2\mathbf{p}_u^T(\mathbf{q}_j-\mathbf{q}_i))\\
&+\beta_1 n||\mathbf{q}_i||^2-2\beta_2 n\mathbf{q}_i^T\mathbf{y}_i\\
\frac{\partial\mathcal{L}}{\partial\mathbf{q}_i}=&\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(2\mathbf{p}_u^T(\mathbf{q}_i-\mathbf{q}_j)\mathbf{p}_u-2\mathbf{p}_u)\\
&+\sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})(2\mathbf{p}_u^T(\mathbf{q}_j-\mathbf{q}_i)(-\mathbf{p}_u)+2\mathbf{p}_u)\\
&+2\beta_1 n\mathbf{q}_i-2\beta_2 n\mathbf{y}_i\\
=&0
\end{aligned}$$

  化简得：

$$\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(\mathbf{p}_u^T\mathbf{q}_i\mathbf{p}_u)+\sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})(-\mathbf{p}_u^T\mathbf{q}_i\mathbf{p}_u)+\beta_1 n\mathbf{q}_i$$
$$=\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(1+\mathbf{p}_u^T\mathbf{q}_j)\mathbf{p}_u+\sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})(\mathbf{p}_u^T\mathbf{q}_i-1)\mathbf{p}_u+\beta_2 n\mathbf{y}_i$$
$$\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(\mathbf{p}_u\mathbf{p}_u^T)\mathbf{q}_i+\sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})(-\mathbf{p}_u\mathbf{p}_u^T)\mathbf{q}_i+\beta_1 n\mathbf{I}_r\mathbf{q}_i$$
$$=\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(1+\mathbf{p}_u^T\mathbf{q}_j)\mathbf{p}_u+\sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})(\mathbf{p}_u^T\mathbf{q}_i-1)\mathbf{p}_u+\beta_2 n\mathbf{y}_i$$

$$\mathbf{q}_i^* = \left(\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(\mathbf{p}_u\mathbf{p}_u^T) + \sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})(-\mathbf{p}_u\mathbf{p}_u^T) + \beta_1 n\mathbf{I}_r\right)^{-1}$$
$$\left(\sum_{u,j} z_u^+ z_u^- r_{ui}(1-r_{uj})(1+\mathbf{p}_u^T\mathbf{q}_j)\mathbf{p}_u + \sum_{u,j} z_u^+ z_u^- r_{uj}(1-r_{ui})(\mathbf{p}_u^T\mathbf{q}_i - 1)\mathbf{p}_u + \beta_2 n\mathbf{y}_i\right)$$

- 运行结果