# 1 Cone programming for A-optimal

The A-optimal is cast as the following SDP.

$$\min \sum_{i=1}^{K} u_i \tag{1}$$

subject to

$$n_{ij} \geq 0 \ \forall\{(i,j)|1 \leq i \leq j \leq K\}, \tag{2}$$

$$\sum_{i} n_{ii} + \sum_{i<j} n_{ij} = N, \tag{3}$$

and

$$\begin{pmatrix} F(\{n_{ab}\}) & \vec{e}_i \\ \vec{e}_i^t & u_i \end{pmatrix} \geq 0 \text{ for } i = 1, 2, \ldots, K \tag{4}$$

where

$$F(\{n_{ab}\}) = \sum_{a=1}^{K} V_{aa} n_{aa} + \sum_{a<b} V_{ab} n_{ab} \tag{5}$$

with sparse matrices $V_{aa}$ and $V_{ab}$ defined by

$$
\begin{aligned}
V_{aa;\alpha\beta} &= s_{aa}^{-2} \delta_{a\alpha} \delta_{a\beta} \\
V_{a\neq b;\alpha\beta} &= s_{ab}^{-2} (\delta_{a\alpha}\delta_{a\beta} + \delta_{b\alpha}\delta_{b\beta} - \delta_{a\alpha}\delta_{b\beta} - \delta_{a\beta}\delta_{b\alpha})
\end{aligned}
\tag{6}
$$

In the notation of the cone programming, there are $K(K+1)/2+K$ variables for minimization:

$$
\begin{aligned}
\vec{x} &= (\vec{n}^t; \vec{u}^t)^t \\
&= (n_{11}, n_{12}, \ldots, n_{1K}, n_{22}, n_{23}, \ldots, n_{K-1\,K}, n_{KK}, u_1, u_2, \ldots, u_K)^t
\end{aligned}
\tag{7}
$$

The coefficients are

$$\vec{c} = (0, 0, \ldots, 0, 1, 1, \ldots, 1)^t \tag{8}$$

with $K(K+1)/2$ entries of 0s and $K$ entries of 1s.

$$A = (1, 1, \ldots, 1, 0, 0, \ldots, 0) \tag{9}$$

with $K(K+1)/2$ entries of 1s and $K$ entries of 0s.

$$
G = - \begin{pmatrix}
& I_{K(K+1)/2} & & & 0_{K(K+1)/2,K} & & \\
\text{vec}(V_{11}^+) & \text{vec}(V_{12}^+) & \cdots \text{vec}(V_{KK}^+) & \text{vec}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & 0 & & \cdots \\
\text{vec}(V_{11}^+) & \text{vec}(V_{12}^+) & \cdots \text{vec}(V_{KK}^+) & 0 & \text{vec}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & \cdots \\
& & & \cdots & & \\
\text{vec}(V_{11}^+) & \text{vec}(V_{12}^+) & \cdots \text{vec}(V_{KK}^+) & 0 & & \cdots & \text{vec}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}
\end{pmatrix}
\tag{10}
$$

where

$$V_{ij}^{+} = \begin{pmatrix} V_{ij} & 0 \\ 0 & 0 \end{pmatrix} \tag{11}$$

$$h = \begin{pmatrix} \vec{0}_{K(K+1)/2} \\ \text{vec}\begin{pmatrix} F(\{n^{(0)}\}) & \vec{e}_1 \\ \vec{e}_1^t & 0 \end{pmatrix} \\ \text{vec}\begin{pmatrix} F(\{n^{(0)}\}) & \vec{e}_2 \\ \vec{e}_2^t & 0 \end{pmatrix} \\ \vdots \\ \text{vec}\begin{pmatrix} F(\{n^{(0)}\}) & \vec{e}_K \\ \vec{e}_K^t & 0 \end{pmatrix} \end{pmatrix} \tag{12}$$

The dimensions of the cone programming is

$$\text{dims} = \begin{cases} \text{'l'}: & K(K+1)/2 \\ \text{'q'}: & [\,] \\ \text{'s'}: & [K+1, K+1, \ldots, K+1] \end{cases} \tag{13}$$

$G$ is a VERY sparse matrix of $K(K+1)/2 + K(K+1)^2$ rows and $K(K+1)/2 + K$ columns, with highly structured entries. Solving the cone programming for large $K$ requires exploiting the structure in the problem.

We need to solve the KKT equations

$$\begin{pmatrix} 0 & A^t & G^t \\ A & 0 & 0 \\ G & 0 & -W^t W \end{pmatrix} \begin{pmatrix} \vec{x} \\ y \\ \vec{z} \end{pmatrix} = \begin{pmatrix} \vec{p} \\ q \\ \vec{l} \end{pmatrix} \tag{14}$$

$\vec{x}$ and $\vec{p}$ are vectors of length $K(K+1)/2 + K$. We will write

$$\begin{aligned} \vec{x} &= (n_{11}, n_{12}, \ldots, n_{KK}; u_1, u_2, \ldots, u_K)^t \\ \vec{p} &= (p_{11}, p_{12}, \ldots, p_{KK}; \pi_1, \pi_2, \ldots, \pi_K)^t \end{aligned} \tag{15}$$

$y$ and $q$ are two scalar numbers.
$\vec{z}$ and $\vec{l}$ are vectors of length $K(K+1)/2 + K(K+1)^2$. We denote

$$\begin{aligned} \vec{z} &= (z_{11}, z_{12}, \ldots, z_{KK}, \text{vec}(\tilde{z}_1)^t, \text{vec}(\tilde{z}_2)^t, \ldots, \text{vec}(\tilde{z}_K)^t)^t \\ \vec{l} &= (l_{11}, l_{12}, \ldots, l_{KK}, \text{vec}(\tilde{l}_1)^t, \text{vec}(\tilde{l}_2)^t, \ldots, \text{vec}(\tilde{l}_K)^t)^t \end{aligned} \tag{16}$$

where each $\tilde{z}_i$ and $\tilde{l}_i$ is a $K+1$ by $K+1$ square symmetric matrix. We will write $\tilde{z}_i$ in a block form

$$\tilde{z}_i = \begin{pmatrix} \tilde{z}_i^* & \zeta_i \\ \zeta_i^t & z_{i,K+1,K+1} \end{pmatrix} \tag{17}$$

$W$ is a block diagonal matrix, satisfying

$$
W^t W \vec{z} = \begin{pmatrix} d_{11}^2 z_{11} \\ d_{12}^2 z_{12} \\ \vdots \\ d_{KK}^2 z_{KK} \\ \text{vec}(r_1 r_1^t \tilde{z}_1 r_1 r_1^t) \\ \text{vec}(r_2 r_2^t \tilde{z}_2 r_2 r_2^t) \\ \vdots \\ \text{vec}(r_K r_K^t \tilde{z}_K r_K r_K^t) \end{pmatrix} \tag{18}
$$

The first group of KKT equations, $A^t y + G^t \vec{x} = \vec{p}$, are

$$
y - z_{ab} - \text{vec}(V_{ab})^t \cdot \text{vec}(\sum_i \tilde{z}_i^*) = p_{ab} \text{ for } 1 \le a \le b \le K \tag{19}
$$

and

$$
-z_{i,K+1,K+1} = \pi_i \text{ for } i = 1, 2, \dots, K \tag{20}
$$

There is only one equation for the second group $A\vec{x} = q$:

$$
\sum_a n_{aa} + \sum_{a<b} n_{ab} = q \tag{21}
$$

The third group of KKT equations, $G\vec{x} - W^t W \vec{z} = \vec{s}$, are

$$
-n_{ab} - d_{ab}^2 z_{ab} = l_{ab} \text{ for } 1 \le a \le b \le K \tag{22}
$$

and

$$
-\begin{pmatrix} F(\{n_{ab}\}) & 0 \\ 0 & u_i \end{pmatrix} - r_i r_i^t \tilde{z}_i r_i r_i^t = \tilde{l}_i \tag{23}
$$

Eq. 23 represents $K(K+1)^2$ equations.

We will first eliminate the variables in $\tilde{z}_i^*$. Denoting

$$
R_i = r_i^{-t} r_i^{-1} = \begin{pmatrix} R_i^* & \gamma_i \\ \gamma_i^t & R_{i,K+1,K+1} \end{pmatrix} \tag{24}
$$

Eq. 23 can be rewritten as

$$
R_i \begin{pmatrix} F(\{n_{ab}\}) & 0 \\ 0 & u_i \end{pmatrix} R_i + \tilde{z}_i = -R_i \tilde{l}_i R_i \tag{25}
$$

Writing $L_i \equiv R_i \tilde{l}_i R_i$ in block form

$$
L_i = \begin{pmatrix} L_i^* & \lambda_i \\ \lambda_i^t & L_{i,K+1,K+1} \end{pmatrix} \tag{26}
$$

Eq. 25 becomes

$$
R_i^* F(\{n_{ab}\}) R_i^* + \gamma_i \gamma_i^t u_i + \tilde{z}_i^* = -L_i^* \tag{27}
$$

$$R_i^* F(\{n_{ab}\})\gamma_i + \gamma_i R_{i,K+1,K+1} u_i + \zeta_i = -\lambda_i \tag{28}$$

and

$$\gamma_i^t F(\{n_{ab}\})\gamma_i + R_{i,K+1,K+1}^2 u_i + z_{i,K+1,K+1} = -L_{i,K+1,K+1} \tag{29}$$

Because $z_{i,K+1,K+1} = -\pi_i$, the last equation is

$$\gamma_i^t F(\{n_{ab}\})\gamma_i + R_{i,K+1,K+1}^2 u_i = \pi_i - L_{i,K+1,K+1} \tag{30}$$

Eq. 27 and Eq. 19 together lead to

$$z_{\alpha\beta} = y - p_{\alpha\beta} + \text{vec}(V_{\alpha\beta})^t \cdot \text{vec}\left(\sum_i R_i^* F(\{n_{ab}\})R_i^* + \sum_i \gamma_i \gamma_i^t u_i\right) + \text{vec}(V_{\alpha\beta})^t \cdot \text{vec}\left(\sum_i L_i^*\right) \tag{31}$$

Plugging the above in Eq. 22, we also eliminate $z_{ab}$.

$$d_{\alpha\beta}^{-2} n_{\alpha\beta} + \text{vec}(V_{\alpha\beta})^t \cdot \text{vec}\left(\sum_i R_i^* F(\{n_{ab}\})R_i^* + \sum_i \gamma_i \gamma_i^t u_i\right) + y$$

$$= p_{\alpha\beta} - \text{vec}(V_{\alpha\beta})^t \cdot \text{vec}\left(\sum_i L_i^*\right) - d_{\alpha\beta}^{-2} l_{\alpha\beta} \tag{32}$$

Eq. 32 ($K(K+1)/2$ equations), Eq. 21 (1 equation), and Eq. 30 ($K$ equations) together determine the variables $\{n_{ab}\}$ ($K(K+1)/2$ variables), $\{u_i\}$ ($K$ variables), and $y$ (1 variable).

The matrix of the form $RF(\{n_{ab}\})R$ is

$$\begin{aligned}
RF(\{n_{ab}\})R &= \sum_a RV_{aa}R n_{aa} + \sum_{a<b} RV_{ab}R n_{ab} \\
&= \sum_a s_{aa}^{-2} R_a R_a^t n_{aa} + \sum_{a<b} s_{ab}^{-2}(R_a R_a^t + R_b R_b^t - R_a R_b^t - R_b R_a^t) n_{ab}
\end{aligned} \tag{33}$$

where $R_a$ is the column vector of the $a$th column of $R$.

The vector of the form $RF(\{n_{ab}\})\gamma$ is

$$RF(\{n_{ab}\})\gamma = \sum_a s_{aa}^{-2} R_a \gamma_a n_{aa} + \sum_{a<b} s_{ab}^{-2}(R_a \gamma_a + R_b \gamma_b - R_a \gamma_b - R_b \gamma_a) n_{ab} \tag{34}$$

The scalar of the form $\gamma^t F(\{n_{ab}\})$ is

$$\gamma^t F(\{n_{ab}\})\gamma = \sum_a s_{aa}^{-2} \gamma_a^2 n_{aa} + \sum_{a<b} s_{ab}^{-2}(\gamma_a^2 + \gamma_b^2 - 2\gamma_a \gamma_b) n_{ab} \tag{35}$$

The inner product $\text{vec}(V_{\alpha\beta})^t \cdot \text{vec}(X)$ is

$$\text{vec}(V_{\alpha\beta})^t \cdot \text{vec}(X) = \begin{cases} s_{\alpha\alpha}^{-2} X_{\alpha\alpha} & \text{if } \alpha = \beta \\ s_{\alpha\beta}^{-2}(X_{\alpha\alpha} + X_{\beta\beta} - X_{\alpha\beta} - X_{\beta\alpha}) & \text{if } \alpha \neq \beta \end{cases} \tag{36}$$

The equation for $\{n_{ab}\}$, $\{u_i\}$, and $y$, summarizing Eq. 32, 21, and 30, is

$$\begin{pmatrix} B & \vec{\eta} \\ \vec{\eta}^t & 0 \end{pmatrix} \begin{pmatrix} \vec{x} \\ y \end{pmatrix} = \begin{pmatrix} \vec{x}_0 \\ y_0 \end{pmatrix} \tag{37}$$

where $\vec{x} = (n_{11}, n_{12}, \dots, n_{22}, n_{23}, \dots, n_{KK}, u_1, u_2, \dots u_K)^t$, and $B$ is a symmetric matrix

$$B = \begin{pmatrix} A & \vec{\eta} \\ \vec{\eta}^t & \mathrm{diag}(R_{i,K+1,K+1}^2) \end{pmatrix} \tag{38}$$

where $\vec{\eta} = (1, 1, \dots 1, 0, 0, \dots 0)^t$ is $K(K+1)/2$ 1's followed by $K$ 0's.

Eq. 37 can be solved by

$$y = \frac{\vec{\eta}^t \cdot B^{-1} \cdot \vec{x}_0 - y_0}{\vec{\eta}^t \cdot B^{-1} \vec{\eta}}$$

$$\vec{x} = B^{-1}\vec{x}_0 - yB^{-1}\vec{\eta} \tag{39}$$

Some algebra shows that the first $K(K+1)/2$-by-$K(K+1)/2$ submatrix of $A$ is given by

$$A_{\alpha\beta,ab} = s_{ab}^{-2}s_{\alpha\beta}^{-2} \cdot \begin{cases} \sum_i R_{i,a,\alpha}^2 & \text{if } a = b, \alpha = \beta \\ \sum_i (R_{i,a,\alpha} - R_{i,b,\alpha})^2 & \text{if } a \neq b, \alpha = \beta \\ \sum_i (R_{i,a,\alpha} - R_{i,a,\beta})^2 & \text{if } a = b, \alpha \neq \beta \\ \sum_i ((R_{i,a,\alpha} - R_{i,b,\alpha}) - (R_{i,a,\beta} - R_{i,b,\beta}))^2 & \text{if } a \neq b, \alpha \neq \beta \end{cases} \tag{40}$$

The computation of this submatrix $A_{\alpha\beta,ab}$ is $O(K^5)$ in time complexity and $O(K^4)$ in memory complexity; it is the most intensive computation in computing the A-optimal by cone programming.

## 2 Including reference input

Often, some of the quantities, say $\{x_a | a \in Q\}$, are measured by other experiments to within certain statistical error $\{\tilde{x}_a \pm \delta_a | a \in Q\}$. The likelihood that our new measurements produce the set of values $\{\hat{x}_e\}$ and the other experiments produce the set of values $\{\tilde{x}_a | a \in Q\}$, given the true values of the quantities of $\{x_i\}$, is

$$\begin{aligned} L &= \prod_i (\sqrt{2\pi}\sigma_i)^{-1} \exp\left(-\frac{(x_i - \hat{x}_i)^2}{2\sigma_i^2}\right) \\ &\cdot \prod_{a \in Q} (\sqrt{2\pi}\delta_a)^{-1} \exp\left(-\frac{(x_a - \tilde{x}_a)^2}{2\delta_a^2}\right) \\ &\cdot \prod_{i<j} (\sqrt{2\pi}\sigma_{ij})^{-1} \exp\left(-\frac{(x_i - x_j - \hat{x}_{ij})^2}{2\sigma_{ij}^2}\right) \end{aligned} \tag{41}$$

Maximizing $\ln L$ with respect to $\{x_i\}$ yields

$$F\vec{x} = \vec{z} \tag{42}$$

where

$$F_{ij} = \begin{cases} \delta_i^{-2} + \sigma_i^{-2} + \sum_{k \neq i} \sigma_{ik}^{-2} & \text{if } i = j \text{ and } i \in Q \\ \sigma_i^{-2} + \sum_{k \neq i} \sigma_{ki}^{-2} & \text{if } i = j \text{ and } i \notin Q \\ -\sigma_{ij}^{-2} & \text{if } i \neq j \end{cases} \tag{43}$$

and

$$z_i = \begin{cases} \sigma_i^{-2} \hat{x}_i + \delta_i^{-2} \tilde{x}_i + \sum_{j \neq i} \sigma_{ij}^{-2} \hat{x}_{ij} & \text{if } i \in Q \\ \sigma_i^{-2} \hat{x}_i + \sum_{j \neq i} \sigma_{ij}^{-2} & \text{if } i \notin Q \end{cases} \tag{44}$$

The A-optimal can be determined by the same SDP of Eqs. 1 to 4, but the Fisher information matrix $F$ now includes diagonal terms of $\delta_i^{-2}$.