A note on Dinkelbach

Frank Wang

Let $\mathbb{E}$ be a finite dimensional Euclidean space; let $\Omega \subseteq \mathbb{E}$ be a convex set, and let $h : \mathbb{E} \to \mathbb{R}$ be sufficiently differentiable. Then, h is called *pseudoconvex* on Ω if

$$x, y \in \Omega, h(y) < h(x) \implies \langle \nabla h(x), y - x \rangle < 0$$

Let $n \geq 3$ and $C, G \in \Omega := \mathbb{S}^n_{++}$ and define

$$f(X) := \mathrm{Tr}(CX), \quad g(X) := \det(X)^{1/n}, \qquad \forall X \in \mathbb{S}^n_{++}$$

Define the fractional function $h(X) := \frac{f(X)}{g(X)}, \quad \forall X \in \mathbb{S}^n_{++}$.

Since f is convex and g is concave, h is pseudoconvex on $\mathbb{S}^n_{++}$.

Motivation for the Dinkelbach Algorithm

Note that

$$f(X) = \mathrm{Tr}(CX) = \mathrm{Tr}(X^{1/2}C^{1/2}C^{1/2}X^{1/2}) = \|C^{1/2}X^{1/2}\|_F^2 > 0$$

since $X, C \succ 0$. Let $\bar{X}$ be the minimizer of $k_{\bar{\lambda}}(X)$ on $\{X \in \Omega : \mathrm{Tr}(X) = 1\}$ and $k_{\bar{\lambda}}(\bar{X}) = 0$. Then,

$$f(\bar{X}) - \bar{\lambda} g(\bar{X}) = 0 \iff \bar{\lambda} = h(\bar{X})$$

For any other $X \in \Omega$ with $\mathrm{Tr}(X) = 1$,

$$0 = k_{\bar{\lambda}}(\bar{X}) \leq k_{\bar{\lambda}}(X) = f(X) - \bar{\lambda} g(X)$$

If $g(X) \leq 0$, then since $\bar{\lambda} \geq 0$,

$$0 < f(X) = \bar{\lambda} g(X) \leq 0$$

which is a contradiction. Otherwise, $g(X) > 0$, so then

$$h(\bar{X}) = \bar{\lambda} \leq \frac{f(X)}{g(X)} = h(X)$$

hence $\bar{X}$ minimizes h subject to $\mathrm{Tr}(X) = 1$, $X \in \Omega$. Thus, if there exists $\bar{\lambda} \geq 0$ such that

$$\bar{X} \in \arg\min\{k_{\bar{\lambda}}(X) : \mathrm{Tr}(X) = 1, X \in \Omega\}, \quad k_{\bar{\lambda}}(\bar{X}) = 0$$

then $\bar{X}$ solves

$$\text{(FP)} \qquad \min h(X) \quad : \quad \mathrm{Tr}(X) = 1, X \in \mathbb{S}^n_{++}$$

The Dinkelbach Algorithm.

(a) **Initialization:** Choose an initial value for the parameter $\lambda_0 \geq 0$

(b) **Iterative Step:** At each iteration k, solve the parametric non-fractional subproblem:

$$\min\{f(X) - \lambda_k g(X) : \mathrm{Tr}(X) = 1, X \in \Omega\}$$

Let X_k be the optimal solution and $k_{\lambda_k}(X_k) = f(X_k) - \lambda_k g(X_k)$ be the optimal value.

(c) **Update step:**

i. If $k_{\lambda_k}(X_k) = 0$, then X_k is an optimal solution to the original fraction problem.

ii. If $k_{\lambda_k}(X_k) \neq 0$, update the parameter $\lambda_{k+1} = \frac{f(X_k)}{g(X_k)}$

(d) **Repeat** the iterative process.

The Lagrangian is

$$L(X, \mu) = \mathrm{Tr}(CX) - \lambda_k \det(X)^{1/n} + \mu(\mathrm{Tr}(X) - 1)$$

From stationarity of KKT,

$$0 = \nabla_X L(X, \mu) = C - \frac{\lambda_k}{n} \det(X)^{1/n-1} \det(X) X^{-1} + \mu I$$

Let $M_X = \frac{\lambda_k}{n} \det(X)^{1/n}$. Then,

$$(\clubsuit) \qquad X = M_X (C + \mu I)^{-1}$$

where we have an implicit constraint that $C + \mu I \succ 0$, since $M_X > 0$. Now, where c_i are the eigenvalues of C,

$$\det(X) = M_X^n \det((C + \mu I)^{-1}) = M_X^n \prod_{i=1}^n \frac{1}{c_i + \mu}$$

hence

$$(\spadesuit) \qquad M_X = \frac{\lambda_k}{n} \left[M_X^n \prod_{i=1}^n \frac{1}{c_i + \mu} \right]^{1/n} = \frac{\lambda_k M_X}{n} \left[\prod_{i=1}^n \frac{1}{c_i + \mu} \right]^{1/n} \implies \left(\frac{\lambda_k}{n} \right)^n = \prod_{i=1}^n (c_i + \mu)$$

since $M_X = 0 \implies \lambda_k = 0$ which implies this X is optimal already, so we can assume $M_X \neq 0$. Now, use the constraint $\mathrm{Tr}(X) = 1$:

$$(\diamondsuit) \qquad 1 = \mathrm{Tr}(M_X (C + \mu I)^{-1}) = M_X \sum_{i=1}^n \frac{1}{c_i + \mu} \implies M_X = \left(\sum_{i=1}^n \frac{1}{c_i + \mu} \right)^{-1}$$

Then, the process is:

1. Solve for μ in ($\spadesuit$) by using Newton's method.
2. Plug it into ($\diamondsuit$) to solve for M_X
3. Use M_X, μ in ($\clubsuit$) to solve for X.

Let

$$p(\mu) := \prod_{i=1}^n (c_i + \mu)$$

Then,

$$\log(p(\mu)) = \sum_{i=1}^n \log(c_i + \mu)$$

hence

$$\sum_{i=1}^n \frac{1}{c_i + \mu} = \frac{d}{d\mu} \log(p(\mu)) = \frac{p'(\mu)}{p(\mu)}$$

hence

$$p'(\mu) = p(\mu) \sum_{i=1}^n \frac{1}{c_i + \mu}$$

This will be used for Newton's method.

Notes

- The problem of solving (♠) is ill-conditioned. In general, $\lambda_k \to 0$, which means $(\lambda_k/n)^n \to 0$, which means the value of μ will be very close to $-\lambda_{\min}(C)$, meaning $C + \mu I$ is very close to singular.
- I tried this algorithm with $g(X) := \log\det(X)$, but the Dinkelbach algorithm requires the denominator $g(X)$ to be positive. Since the answer often resides near the boundary of the simplex $\text{Tr}(X) = 1$, $\log\det(X)$ is often negative near the optimal value.