Research Progress 7.8

Yang Xiao

1 PREVIOUS PROGRESS

Given PSD matrices A, B, we can expand $per(A+B)$ as:

$$per(A+B) = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}\langle \mathbb{1}, (A^{\otimes k}\otimes B^{\otimes n-k})\mathbb{1}\rangle$$

We can further expand the bilinear term by:

$$\langle \mathbb{1}, (A^{\otimes k}\otimes B^{\otimes n-k})\mathbb{1}\rangle = k!(n-k)! \sum_{S,T\in\binom{[n]}{k}} per(A[S;T])per(B[\bar{S};\bar{T}])$$

So we have:

$$per(A+B) = \sum_{k=0}^{n}\sum_{S,T\in\binom{[n]}{k}} per(A[S;T])per(B[\bar{S};\bar{T}])$$

Therefore, if we can utilize the fact that $\begin{pmatrix} A[S;T] & 0 \\ 0 & B[\bar{S};\bar{T}]\end{pmatrix}$ is low rank and come up with an SDP relaxation to better approximate each bilinear term, we get a better approximation for $per(A+B)$.

Now let's consider the SDP relaxation for $\sum_{S,T\in\binom{[n]}{k}} per(A[S;T])per(B[\bar{S};\bar{T}])$:

$$\begin{aligned} \min \sum_{S\in\binom{n}{k}} \prod_{i=1}^{k} D_{S_i,S_i} \prod_{j=1}^{n-k} Q_{\bar{S}_j,\bar{T}_j} \\ s.t.\ D \succeq A \\ Q \succeq B \end{aligned} \tag{1}$$

We have proven that this convex relaxation is an upper bound of $\sum_{S,T\in\binom{n}{k}} per(A[S;T])per(B[\bar{S};\bar{T}])$, and is also poly-time solvable. Now we are trying to find a metaphor relation between this and some maximum product of linear forms.

2 THIS WEEK'S WORK

2.1 Lagrangian duality

We first show that the following pair of programmings are dual to each other:

$$\begin{aligned} \mu^*(A) := \min \prod_{i=1}^{n} D_{i,i} \\ s.t.\ D \succeq A \end{aligned} \tag{1}$$

$$\begin{aligned}\nu^*(A) := \max \prod_{i=1}^{n} v_i^* P v_i \\ s.t. tr(P) = n \\ D \succeq 0\end{aligned} \tag{2}$$

1. **Approach 1: Taking the dual of $\mu^*(A)$ and successfully end up with $\nu^*(A)$**

   Let $A = V^*V$ Notice that $\mu^*(A)$ is equivalent to the following program:

   $$\begin{aligned}\min \prod_{i=1}^{n} \frac{1}{\alpha_i} \\ s.t.\ Diag^{-1}(\alpha) \succeq V^*V\end{aligned} \tag{3}$$

   which is also equivalent to:

   $$\begin{aligned}\min \prod_{i=1}^{n} \frac{1}{\alpha_i} \\ s.t.\ \sum_{i=1}^{n} \alpha_i v_i v_i^* \preceq I\end{aligned} \tag{4}$$

   We first take a log on the objective function to get the log of dual, at last we eliminate the log. Define:

   $$\begin{aligned}\mu^{**}(A) := \min -\sum_{i=1}^{n} \ln \alpha_i \\ s.t.\ \sum_{i=1}^{n} \alpha_i v_i v_i^* \preceq I\end{aligned} \tag{5}$$

   Let $P \succeq 0$ be the Lagrangian multiplier for the only inequality, then the Lagrangian function is:

   $$L(\alpha, P) = -\sum_{i=1}^{n} \ln \alpha_i - < P, \sum_{i=1}^{n} \alpha_i v_i v_i^* - I >$$

   Then the dual of $\mu^{**}(A)$ is $\sup_P \inf_\alpha L(\alpha, P)$. Now fix P, $L(\alpha, P)$ is minimized when the gradient of α is 0, i.e.

   $$\nabla_\alpha L = (-\frac{1}{\alpha_1} + < P, v_1 v_1^* >, ..., -\frac{1}{\alpha_n} + < P, v_n v_n^* >)^T = 0$$
   $$\Rightarrow \alpha_i = \frac{1}{< P, v_i v_i^* >}$$

   Hence, the dual of $\mu^{**}(A)$ is:

   $$\sup_P \sum_{i=1}^{n} \ln < P, v_i v_i^* > + n - tr(P)$$

Further, since there exists slater point in primal program (take α to be sufficiently small), strong duality holds, and any pair of optimal solution (α^*, P^*) must satisfy KKT condition, in particular, they must satisfy stationarity and complimentary slackness, i.e.

$$\nabla_\alpha L(\alpha^*, P^*) = 0 \Rightarrow \alpha_i^* = \frac{1}{< P^*, v_i v_i^* >} \textbf{ (stationarity)}$$
$$- < P^*, \sum_{i=1}^n \alpha_i^* v_i v_i^* - I >= 0 \Rightarrow tr(P) = n \textbf{ (complementarity)}$$

Therefore, the dual program of $\mu^{**}(A)$ is the following:

$$\begin{aligned} \nu^{**}(A) := \max \sum_{i=1}^n \ln < P, v_i v_i^* > \\ s.t.\ tr(P) = n \\ P \succeq 0 \end{aligned} \tag{6}$$

By exponentiate the objective function, we've recovered $\nu^*(A)$.

2. **Approach 2: Taking the dual of $\nu^*(A)$ but failed**
   By taking the equivalent form and taking a log on the objective function, we get the following program:

$$\begin{aligned} \nu^{**}(A) := \max \ln \prod_{i=1}^n t_i \\ s.t.\ < v_i v_i^*, P >\geq t_i\ \forall i \in [n] \\ < I, P >= n \\ P \succeq 0 \end{aligned} \tag{7}$$

   which is equivalent to:

$$\begin{aligned} \nu^{**}(A) := -\min - \ln \prod_{i=1}^n t_i \\ s.t.\ < v_i v_i^*, P >\geq t_i\ \forall i \in [n] \\ < I, P >= n \\ P \succeq 0 \end{aligned} \tag{7}$$

We first ignore the negative sign before 'min' on the objective function, and try to add it back after we get the dual. Consider the Lagrangian function:

$$\begin{aligned} L(P, t, \alpha, \lambda, Q) &= -\ln \prod_{i=1}^n t_i - \sum_{i=1}^n \alpha_i (< v_i v_i^*, P > - t_i) + \lambda(< I, P > - n) - < Q, P > \\ &= -\ln \prod_{i=1}^n t_i + \sum_{i=1}^n \alpha_i t_i - \lambda n + < \lambda I - Q - \sum_{i=1}^n \alpha_i v_i v_i^*, P > \end{aligned}$$

where $\alpha_i \geq 0$, λ (free), $Q \succeq 0$ are Lagrangian multipliers. Then the dual program is

$$\sup_{\alpha \geq 0, \lambda, Q \succeq 0} \inf_{P, t > 0} L(P, t, \alpha, \lambda, Q)$$

Note that fixing α, λ, Q, we can always take $P = -k(\lambda I - Q - \sum_{i=1}^{n} \alpha_i v_i v_i^*)$, where k is a positive constant. If $\lambda I - Q - \sum_{i=1}^{n} \alpha_i v_i v_i^* \neq 0$, then $L(P, t, \alpha, \lambda, Q) \to -\infty$ as $k \to \infty$, so $\inf_{P,t>0} L(P, t, \alpha, \lambda, Q) = -\infty$, which will not be an optimal solution for dual.

Thus, the dual program is equivalent to:

$$\sup_{\alpha \geq 0, \lambda, Q \succeq 0} \inf_{P, t > 0} -\ln \prod_{i=1}^{n} t_i + \sum_{i=1}^{n} \alpha_i t_i - \lambda n$$

$$s.t. \ \lambda I - Q - \sum_{i=1}^{n} \alpha_i v_i v_i^* = 0 \Leftrightarrow \lambda I \succeq \sum_{i=1}^{n} \alpha_i v_i v_i^*$$

Following the same procedure as Approach 1, fix feasible λ, α, Q that satisfies $\lambda I \succeq \sum_{i=1}^{n} \alpha_i v_i v_i^*$, then the Lagrangian function $L(P, t, \alpha, \lambda, Q)$ is minimized when the gradient of t is 0, i.e.

$$\nabla_t L = (-\frac{1}{t_1} + \alpha_1, ..., -\frac{1}{t_n} + \alpha_n)^T = 0 \Rightarrow t_i = \frac{1}{\alpha_i}$$

So the dual program is equivalent to:

$$\sup_{\alpha \geq 0, \lambda, Q \succeq 0} \ln \prod_{i=1}^{n} \alpha_i - (\lambda - 1) n$$

$$s.t. \ \lambda I - Q - \sum_{i=1}^{n} \alpha_i v_i v_i^* = 0 \Leftrightarrow \lambda I \succeq \sum_{i=1}^{n} \alpha_i v_i v_i^*$$

And thus I'm stuck here and don't know what to do.

However, if instead of taking the gradient of α to be 0, we take $\alpha_i = \frac{\lambda}{t_i}$, then the "fake" dual would be:

$$\sup_{\alpha \geq 0, \lambda, Q \succeq 0} \ln \prod_{i=1}^{n} \frac{\lambda}{\alpha_i}$$

$$s.t. \ \lambda I - Q - \sum_{i=1}^{n} \alpha_i v_i v_i^* = 0 \Leftrightarrow \lambda I \succeq \sum_{i=1}^{n} \alpha_i v_i v_i^*$$

by taking exponential on the objective function and make a change of variable, this program is equivalent to $\mu^*(A)$, which I'm still not clear why.