Exercise session notes - Week 12

Formulations as SDP + Schur Complement

Consider a quadratic program of the form

$$\begin{aligned} \min \quad & c^\top x \\ x^\top Q x + b^\top x + d &\leq 0 \end{aligned}$$

If Q is positive semi-definite, then this problem is convex and we can solve it efficiently via Newton Descent. But what if Q has negative eigenvalues? Then the problem is no longer convex, but we can still solve it partially with a relaxation to a SDP. We first define the matrix $X = xx^\top$ which is a symmetric matrix and contains $x_i x_j$ on the ith row and jth column. This is often used when we have a quadratic form because then

$$x^\top Q x = \sum_{i,j} x_i Q_{ij} x_j = \sum_{i,j} X_{ij} Q_{ij} = \mathrm{Tr}(XQ) = \langle X, Q \rangle$$

and this is an affine function on S_+^n. Therefore our QP is equivalent to

$$\begin{aligned} \min \quad & c^\top x \\ \langle X, Q \rangle + b^\top x + d &\leq 0 \\ X &= xx^\top \end{aligned}$$

Finally we can relax the last constraint and we get the SDP

$$\begin{aligned} \min \quad & c^\top x \\ \langle X, Q \rangle + b^\top x + d &\leq 0 \\ X &\succeq 0 \end{aligned}$$

Note that by solving this problem we obtain a lower bound on the optimal value of the QP.

A very important fact on positive semi-definite matrices is the Schur Complement:

Proposition (Schur Complement). *Let X be a symmetric block matrix*

$$X = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix}$$

If $A \succ 0$ then

$$X \succeq 0 \iff C - B^\top A^{-1} B \succeq 0$$

If $C \succ 0$ then

$$X \succeq 0 \iff A - B C^{-1} B^\top \succeq 0$$

Consider the following 0-1 QP

$$\begin{aligned} \max \quad & y^\top B y \\ y &\in \{0,1\}^n \end{aligned}$$

We define as before $Y = yy^\top$ and note that

$$y_i^2 = y_i \implies \operatorname{diag}(Y) = y$$

In particular we can write

$$Y = \operatorname{diag}(Y)\operatorname{diag}(Y)^\top$$

We can relax this constraint into the following

$$Y - \operatorname{diag}(Y)\operatorname{diag}(Y)^\top \succeq 0$$

and by the Schur Complement we can write a relaxation of the QP program as the SDP

$$\begin{aligned} \max \quad & \langle Y, B \rangle \\ \begin{pmatrix} Y & \operatorname{diag}(Y) \\ \operatorname{diag}(Y)^\top & 1 \end{pmatrix} & \succeq 0 \\ Y & \succeq 0 \end{aligned}$$

As before, by solving this problem we obtain a lower bound on the optimal value of the 0-1 QP.

Another application of the Schur Complement comes from geometry: We are given k points $z_1, \ldots, z_k \in \mathbb{R}^n$ and we want to find an Ellipse E that contains all the points and has minimal volume. We can describe an Ellipse as

$$E_{P,c} = \left\{ x \;\middle|\; (x-c)^\top P(x-c) \right\}$$

where $c \in \mathbb{R}^n$ is the center of the Ellipse and $P \succ 0$ is the matrix that gives us the stretch of the Ellipse (if $P = I$ then we obtain a circle). Moreover the volume of the Ellipse is proportional to

$$\operatorname{vol}(E_{P,c}) \sim \sqrt{\det(P^{-1})}$$

Therefore we can model this problem as the following abstract program

$$\begin{aligned} \min_{P,c} \quad & \operatorname{vol}(E_{P,c}) \sim \sqrt{\det(P^{-1})} \\ \text{s.t.} \quad & z_i \in E_{P,c} \qquad \forall i \end{aligned}$$

Since the function $\log(\cdot)$ is monotone, we can apply that to the objective function and the problem remains the same.

$$\begin{aligned} \min_{P,c} \quad & \log\left(\sqrt{\det(P^{-1})}\right) = -\frac{1}{2}\log\det(P) \\ \text{s.t.} \quad & (z_i - c)^\top P(z_i - c) \le 1 \qquad \forall i \end{aligned}$$

Recall that $\log\det(\cdot)$ is a concave function. Since $P \succ 0$ we can write it as $P = B^\top B$ and we get

$$\begin{aligned} \min_{B,c} \quad & -\frac{1}{2}\log\det(B^\top B) = -\log\det(B) \\ \text{s.t.} \quad & 1 - (z_i - c)^\top B^\top B(z_i - c) \ge 0 \qquad \forall i \end{aligned}$$

Finally, we can define $d := Bc$ and apply the Schur Complement to the constraint so that obtain the SDP

$$\begin{aligned} \min_{B,d} \quad & -\log\det(B) \\ \text{s.t.} \quad & \begin{pmatrix} I & Bz_i - d \\ (Bz_i - d)^\top & 1 \end{pmatrix} \succeq 0 \qquad \forall i \end{aligned}$$

After we have solved this program we can obtain a solution to the original problem by computing $P = B^\top B$ and $c = B^{-1}d$.