Assignment 5, Optimization methods (Maxim Kaledin)

Problem 3

Apply Sylvester's criterium to ensure that matrix is positive semidefinite. All principal minors should be non-negative, these are

$$\begin{aligned}
-25y_1^2 &\geq 0, \\
(4y_1+2)(-25y_1^2) &\geq 0, \\
3y_2(4y_1+2) &\geq 0, \\
3y_2 &\geq 0, \\
4y_1+2 &\geq 0.
\end{aligned}$$

The first inequality gives $y_1 = 0$, and as result constraints are $D = \{y : y_1 = 0, y_2 \geq 0\}$. Obviously, all feasible solutions are optimal and optimal goal is 0.

Derive the dual problem (after we dismiss redundant constraints):

$$\max_{\lambda,\nu} \inf_{y \in D} \left[2y_1 + \nu y_1 - \lambda y_2\right], \quad \text{s.t. } \lambda \geq 0.$$

Since $y \in D$, $\inf_{y\in D}\left[2y_1 + \nu y_1 - \lambda y_2\right] = \inf_{y_2 \in D}\left[-\lambda y_2\right]$ and

$$\inf_{y_2 \in D}\left[-\lambda y_2\right] = \begin{cases} -\infty, & \lambda \geq 0, \\ 0, & \lambda = 0. \end{cases}$$

So $\lambda = 0$ is the solution to dual problem with optimal goal equals 0.

Problem 4

1

The problem can be reformulated as

$$\begin{aligned}
&\min_{x,t} t \quad s.t. \\
&t \geq a_i^T x + b_i \quad \forall i \in \{1,..,m\}.
\end{aligned}$$

Corresponding Lagrange dual problem is

$$\begin{aligned}
&\max_{\lambda} \inf_{x,t}\left[t + \sum_{i=1}^{m} \lambda_i (a_i^T x + b_i - t)\right], \\
&s.t. \quad \lambda \geq 0,
\end{aligned}$$

where $\lambda_i \geq 0$ are dual variables.

2

Let

$$\hat{f}_2(y) = \frac{1}{\alpha} \log\left(\sum_{i=1}^{m} \exp \alpha y_i\right),$$

and $y_i = a_i^T x + b_i$ be new constraints. First, formulate the primal problem:

$$\begin{aligned} &\min_{x,y} \hat{f}_2(y) \\ &s.t. \quad y_i = a_i^T x + b_i \quad \forall i \in \{1, .., m\}\,. \end{aligned}$$

The dual one is defined as

$$\max_{\lambda} \inf_{x,t} \left[t + \sum_{i=1}^{m} \lambda_i (a_i^T x + b_i - t) \right]$$

without constraints on λ because the constraints are equalities.

3

Let us show the right side of inequality, in order to do that we can exploit monotonicity of exponent:

$$\begin{aligned} &\frac{1}{\alpha} \log \sum_{i=1}^{m} \exp\left(\alpha(a_i^T x + b_i)\right) - \max_i (a_i^T x + b_i) \leq \\ &\leq \frac{1}{\alpha} \log\left(m \exp\left(\alpha \max_i (a_i^T x + b_i)\right)\right) - \max_i (a_i^T x + b_i) = \\ &= \frac{\log(m)}{\alpha}. \end{aligned}$$

Problem 5

Define Lagrange function of a problem $L(x, \lambda) = \|Ax - b\|_2^2 + \sum_{i=1}^{m} \lambda_i (c_i.x - d_i)$. Since there are no inequality constraints, KKT conditions are exactly

$$\begin{aligned} \frac{\partial L}{\partial x} &= 0 \\ \frac{\partial L}{\partial \lambda} &= 0, \end{aligned}$$

where derivatives denote the gradients truncated to needed set of variables. These sets of constraints are exactly

$$\begin{aligned} &\frac{\partial L}{\partial x_j} = 2 \sum_{i=1}^{m} a_{ij} \sum_{k=1}^{n} (a_{ik} x_k - b_i) + \sum_{i=1}^{p} \lambda_i c_{ij}, \\ &Cx = d. \end{aligned}$$

Let us write the first expression shorter:

$$\nabla_x L = 2A^T (Ax - b) + C^T \lambda.$$

Since the goal function is convex, solution to system

$$\begin{cases} 2A^T(Ax-b)+C^T\lambda=0, \\ Cx=d \end{cases}$$

exists and is an optimal solution of primal problem. With pseudo-inverse (it exists since C has full row rank):

$$\begin{aligned} &x=(C^TC)^{-1}C^Td, \\ &\lambda: \quad C^T\lambda=-2A^T(Ax-b). \end{aligned}$$

Now derive the dual problem:

$$\max_{\lambda}\inf_{x} L(x,\lambda).$$

Lagrangian L remains convex for all λ, so the solution to inner problem is given by the system of normal equations

$$2A^T(Ax-b)+C^T\lambda=0$$

and can be found based on any λ since matrix A^TA has full rank. We could write the solution explicitly:

$$x=(A^TA)^{-1}\left(-\frac{1}{2}C^T\lambda+A^Tb\right).$$

Problem 6

(1)

Consider problem (4) firstly:

$$\begin{aligned} &\min_{x} c^Tx \\ &s.t. \quad Ax=b,\ x\geq 0. \end{aligned}$$

Its dual problem is

$$\begin{aligned} &\max_{\lambda,\mu}\inf_{x}\left(c^Tx+\sum_{i=1}^{m}\lambda_i(a_{i\cdot}x-b_i)-\sum_{j=1}^{n}\mu_jx_j\right), \\ &s.t. \quad \mu\geq 0, \end{aligned}$$

and KKT-conditions are

- $c+A^T\lambda-\mu=0$ (stationarity, $\nabla_x L(x,\lambda,\mu)=0$),
- feasibility conditions (as in problem statement),
- $Ax=b$ (from $\nabla_\lambda L(x,\lambda,\mu)=0$),
- $\mu_jx_j=0 \quad \forall j$ (complementary slackness).

(2)

The idea is the same. Consider problem (5):

$$\min_x c^T x - \tau \sum_{i=1}^{n} \log(x_i)$$
$$s.t. \quad Ax = b.$$

Its dual problem is (let $(1/x)$ denotes coordinate-wise division)

$$\max_\lambda \inf_x \left(c^T x - \tau \sum_{i=1}^{n} \log(x_i) + \sum_{i=1}^{m} \lambda_i (a_{i\cdot} x - b_i) \right).$$

and KKT-conditions are

- $c - \tau(1/x) + A^T \lambda = 0$ (stationarity, $\nabla_x L(x, \lambda, \mu) = 0$),
- feasibility conditions (as in problem statement),
- $Ax = b$ (from $\nabla_\lambda L(x, \lambda, \mu) = 0$).

Note that the goal is well-defined only for non-negative x, otherwise the result is a complex number.