# 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.

Let $F_1, F_2, G$ are such symmetric matrices that the initial matrix $A = y_1F_1 + y_2F_2 + G$:

$$
F_1 = \begin{bmatrix} 0 & 5 & 0\\ 5 & 0 & 0\\ 0 & 0 & 4 \end{bmatrix}, F_2 = \begin{bmatrix} 0 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 0 \end{bmatrix}, G = \begin{bmatrix} 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 2 \end{bmatrix}.
$$

Then we can write down the dual problem in the following way (denoting $Z \succeq 0$ as symmetric matrix of dual variables and scalar product as $(A, B) = \mathrm{tr}(A^TB)$):

$$
\begin{aligned}
&\max_{Z\succeq 0} -(Z,G), \quad s.t.\\
&(Z,F_1) = 2, (Z,F_2) = 0.
\end{aligned}
$$

Rewrite it in the coordinate form:

$$
\begin{aligned}
&\max_{Z\succeq 0} -2z_{33}, \quad s.t.\\
&10z_{12} + 4z_{33} = 2, \qquad\qquad z_{22} = 0.
\end{aligned}
$$

Now add Sylvester's criterium inequalities

$$
\begin{aligned}
&z_{11} \geq 0,\ z_{22} \geq 0,\ z_33 \geq 0,\\
&z_{11}z_{22} - z_{12}^2 \geq 0,\\
&z_{11}z_{33} - z_{13}^2 \geq 0,\\
&z_{22}z_{33} - z_{23}^2 \geq 0\\
&det(Z) \geq 0.
\end{aligned}
$$

This conditions can be significantly simplified if we add constraints from the dual problem. Finally, we obtain

$$
\begin{aligned}
&\max_{Z} -2z_{33}, \quad s.t. \\
&10z_{12} + 4z_{33} = 2, \\
&z_{22} = 0, z_{12} = 0, z_{23} = 0 \\
&z_1 \geq 0, z_{33} \geq 0, \\
&z_{11}z_{33} - z_{13}^2 \geq 0.
\end{aligned}
$$

Since $z_{12} = 0$ the entry $z_{33} = 1/2$ and we can somehow assign other variables to satisfy feasibility. Thus the optimal dual goal is $-1$ which is not equal to optimal primal goal 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. We can easily simplify this because it is a linear program:

$$
\begin{aligned}
&\max_{\lambda} b^T \lambda, \quad s.t. \\
&\begin{bmatrix} A \\ -1 ... -1 \end{bmatrix} \lambda = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix}
\end{aligned}
$$

### 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

$$\max_\lambda \inf_{x,y} \left[\frac{1}{\alpha}\log\left(\sum_{i=1}^m \exp \alpha y_i\right) + \sum_{i=1}^m \lambda_i(a_i^T x + b_i - y_i)\right] =$$
$$= \inf_y \left[\frac{1}{\alpha}\log\left(\sum_{i=1}^m \exp \alpha y_i\right) - \sum_{i=1}^m \lambda_i y_i\right] + \inf_{x,y}\left[\sum_{i=1}^m \lambda_i(a_i^T x + b_i)\right].$$

**3**

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

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

## 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

$$\frac{\partial L}{\partial x} = 0$$
$$\frac{\partial L}{\partial \lambda} = 0,$$

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

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

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):

$$x = (C^T C)^{-1} C^T d,$$
$$\lambda : \quad C^T\lambda = -2A^T(Ax - b).$$

Now derive the dual problem:

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

Lagrangian $L$ remains convex for all $\lambda$, 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 $\lambda$ 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

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

It can be simplified to

$$\begin{aligned} &\max_{\lambda} b^T\lambda, \\ &s.t. \ A^Ty \leq c. \end{aligned}$$

KKT-conditions for primal problem are

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

### (2)

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

$$\begin{aligned} &\min_{x} c^Tx - \tau \sum_{i=1}^{n} \log(x_i) \\ &s.t. \quad Ax = b. \end{aligned}$$

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 $\nabla_\lambda L(x, \lambda, \mu) = 0$).

Since the Hessian matrix of inf sub-problem is positive definite (it is even diagonal), infimum is achieved when neccessary conditions are satisfied (i.e. stationarity):

$$(1/x) = \frac{1}{\tau}(c + A^T\lambda) = y(\lambda).$$

So, simplified dual is

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

After some manipulation we can obtain

$$x_i = \frac{\tau}{c_i + (a_{\cdot i})^T \lambda}$$

and simplify the dual even more

$$\max_{\lambda} \left( \tau \sum_{i=1}^{n} \frac{c_i}{c_i + (a_{\cdot i})^T\lambda} - n\tau \log \tau + \tau \sum_{i=1}^{n} \log(c_i + a_i^T \lambda) + \sum_{i=1}^{m} \lambda_i (a_{i\cdot}(1/y(\lambda)) - b_i) \right).$$

Consider closely the sum

$$\sum_{i=1}^{m} \lambda_i a_{i\cdot}(1/y(\lambda)) = \sum_{i=1}^{m} \lambda_i \sum_{j=1}^{n} \frac{a_{ij}\tau}{c_j + (a_{\cdot j})^T\lambda} =$$
$$= \sum_{j=1}^{n} \sum_{i=1}^{m} \frac{\lambda_i a_{ij} \tau}{c_j + (a_{\cdot j})^T \lambda} = \sum_{j=1}^{n} \frac{(a_{\cdot j})^T \lambda \tau}{c_j + (a_{\cdot j})^T\lambda}.$$

It cancels with the first sum and results in $m\tau$. Finally, the dual problem is

$$\max_{\lambda} \left( \tau \sum_{j=1}^{n} \log(c_j + (a_{\cdot j})^T \lambda) - b^T \lambda \right).$$