Exam

Exercise 1

Let $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{n\times m}$. We want the solution of $\min \|x\|_\infty$ s.t. $Ax = b$ (P)

$$(P) \iff \min_{t,x} \ t \quad \text{s.t.} \ \|x\|_\infty \leq t, \ Ax = b \quad \iff \min_{t,x} \ t \quad \text{s.t.} \ |x_1| \leq t, \dots, |x_n| \leq t, \ Ax = b$$

$$\iff \min_{t,x} \ t \quad \text{s.t.} \ -t \leq x_1 \leq t, \dots, -t \leq x_n \leq t, \ Ax = b \quad \iff \min_{t,x} \ t \quad \text{s.t.} \ x \leq t\mathbb{1}_n, \ x \geq -t\mathbb{1}_n, \ Ax = b$$

Therefore (P) can be rewritten as a LP:

$$\min_{t,x} \ t \quad \text{s.t.} \quad x \leq t\mathbb{1}_n, \quad -x \leq t\mathbb{1}_n, \quad Ax = b \qquad \text{(LP)}$$

The associated Lagrangian is:

$$L(t, x, \lambda_1, \lambda_2, \nu) = t + \lambda_1^T(x - t\mathbb{1}_n) + \lambda_2^T(-x - t\mathbb{1}_n) + \nu^T(Ax - b)$$
$$= t(1 - \lambda_1^T\mathbb{1}_n - \lambda_2^T\mathbb{1}_n) + (\lambda_1^T - \lambda_2^T + \nu^T A)x - \nu^T b$$

The dual function is

$$g(\lambda_1, \lambda_2, \nu) = \begin{cases} -\nu^T b & \text{if } \begin{cases} (\lambda_1 + \lambda_2)^T\mathbb{1}_n = 1 \\ \lambda_1 - \lambda_2 + A^T\nu = 0 \end{cases} \\ -\infty & \text{otherwise} \end{cases}$$

And the dual problem is : $\max \quad g(\lambda_1, \lambda_2, \nu)$

$\text{s.t.} \quad \lambda_1 \succeq 0$

$\lambda_2 \succeq 0$

which is equivalent to $\max \quad -\nu^T b$

$\text{s.t.} \quad \lambda_1 \succeq 0 \ , \ \lambda_2 \succeq 0$

$(\lambda_1 + \lambda_2)^T \mathbb{1}_n = 1$

$\lambda_1 - \lambda_2 + A^T\nu = 0$

We can finally rewrite the dual problem as

$$\begin{aligned} \max \quad & -\nu^T b \\ \text{s.t.} \quad & \lambda \succeq 0 \\ & \lambda + A^T\nu \succeq 0 \\ & (2\lambda + A^T\nu)^T \mathbb{1}_n = 1 \end{aligned}$$

Exercise 2

Let $u, v \in \mathbb{R}^m_{++}$. $D_{KL}(u, v) = \sum_{i=1}^m u_i \log\left(\frac{u_i}{v_i}\right) - u_i + v_i$

If one notes $f(u) = \sum_{i=1}^m u_i \log(u_i)$ the negative entropy of u, one has:

$\nabla f(u) = \sum_{i=1}^m (\log(u_i) + 1) e_i$ on $\mathbb{R}^n_{++}$ where $(e_i)_{i \in \{1,\dots,n\}}$ is the canonical basis of $\mathbb{R}^n$.

Therefore $f(u) - f(v) - \nabla f(v)^T(u - v) =$

$= \sum u_i \log(u_i) - \sum v_i \log(v_i) - \sum (\log(v_i) + 1)(u_i - v_i)$

$= \sum u_i (\log(u_i) - \log(v_i)) - u_i + v_i \quad + 0$

$= D_{KL}(u, v)$

Moreover one has that $h : x \mapsto x \log(x)$ is strictly convex for $x \in \mathbb{R}_{++}$

(because $h''(x) = 1/x > 0$) so f is strictly convex (as a sum of strictly convex functions)

Moreover, f is differentiable and therefore the first order condition of strict convexity gives that $f(u) > f(v) + \nabla f(v)^T(u-v)$

$$\Rightarrow \quad D_{kl}(u,v) = f(u) - f(v) - \nabla f(v)^T(u-v) > 0$$

Finally, if $u = v \in \mathbb{R}^n_{++}$, then $D_{kl}(u,v) = 0$

Hence: $\quad D_{kl}(u,v) \geq 0 \quad \forall u,v \in \mathbb{R}^n_{++}$

$\quad D_{kl}(u,v) = 0 \iff u = v$

Exercise 3

Let $\begin{cases} A \in S_n, \ b \in \mathbb{R}^n, \ c \in \mathbb{R}^n, \ D \in \mathbb{R}^{m \times n}, \ g \in \mathbb{R}^m \\ A \succeq 0 \text{ convex} \end{cases}$

and $\quad (P): \quad \min \quad c^T x$

$$\text{s.t.} \quad x^T(A - bb^T)x \leq 0$$

$$b^T x \geq 0$$

$$Dx = g$$

Since $A \in S_n$ and $A \succeq 0$, there exists $V \in S_n$ such that $A = V^T V$.

Indeed, by the spectral theorem, there exists O orthogonal and Σ diagonal with positive values such that $A = O^T \Sigma O$.

The matrix $V = O^T \Sigma^{1/2} O$ works

Then one can rewrite the condition $x^T(A - bb^T)x \leq 0$

as $x^T V^T V x - x^T b b^T x \leq 0 \iff \|Vx\|_2^2 \leq (b^T x)^2$

Since the condition $b^T x \geq 0$ holds, the above inequality is equivalent to $\|Vx\|_2 \leq b^T x$

Therefore, we can rewrite (P) as a SCOP:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \|Vx\|_2 \leq b^T x \\ & Dx = g \end{aligned}$$

(SOCP) which is a convex problem

Now introduce a new variable $y = Vx$

$$\text{(SOCP)} \iff \begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \|y\|_2 \leq b^T x \\ & Vx = y \\ & Dx = g \end{aligned}$$

The Lagrangian is (for $\lambda \in \mathbb{R}$, $\nu_1 \in \mathbb{R}^n$, $\nu_2 \in \mathbb{R}^m$)

$$L(x, y, \lambda, \nu_1, \nu_2) = c^T x + \lambda(\|y\|_2 - b^T x) + \nu_1^T(Vx - y) + \nu_2^T(Dx - g)$$

$$= (c - b\lambda + V^T \nu_1 + D^T \nu_2)^T x + \lambda \|y\|_2 - \nu_1^T y - \nu_2^T g$$

Now recall that the conjugate of $x \mapsto \|x\|$ is $y \mapsto \begin{cases} 0 & \text{if } \|y\|_* \leq 1 \\ \infty & \text{otherwise} \end{cases}$

and that the ℓ_2 norm is self dual.

Then $\inf_y (\lambda \|y\|_2 - \nu_1^T y) = -\lambda \sup\left(\frac{1}{\lambda}\nu_1^T y - \|y\|_2\right) \quad (\lambda \neq 0)$

$$= \begin{cases} 0 & \text{if } \|\nu_1\|_2 \leq |\lambda| \\ -\infty & \text{otherwise} \end{cases}$$

the obtained formula also works for the case $\lambda = 0$.

Therefore the dual function is:

$$g(\lambda, \nu_1, \nu_2) = \begin{cases} -\nu_2^T g & \text{if } \begin{cases} c - b\lambda + V^T\nu_1 + D^T\nu_2 = 0 \\ \|\nu_1\|_2 \leq |\lambda| \end{cases} \\ -\infty & \text{otherwise} \end{cases}$$

The dual problem can be written as:

$$\begin{aligned} \max \quad & g(\lambda, \nu_1, \nu_2) & \iff \quad \max \quad & -\nu_2^T g \\ \text{s.t.} \quad & \lambda \geq 0 & \text{s.t} \quad & c - b\lambda + V^T\nu_1 + D^T\nu_2 = 0 \\ & & & \|\nu_1\|_2 \leq |\lambda| \\ & & & \lambda \geq 0 \end{aligned}$$

$$\iff \quad \begin{aligned} \max \quad & -\nu_2^T g \\ \text{s.t.} \quad & c - \lambda b + V^T\nu_1 + D^T\nu_2 = 0 \\ & \|\nu_1\|_2 \leq \lambda \end{aligned}$$

Exercise 4

Let (P): $\min \; -\sum_{i=1}^{m} \log(b_i - a_i^T x)$ over $\overbrace{\{x \mid a_i^T x < b_i, \, i = 1, \dots, m\}}^{\mathcal{D}}$

Introduce a new variable $y = b - Ax$ where the ith row of A is a_i^T, so that $y_i = b_i - a_i^T x$.

The new problem is

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} \log(y_i) \\ \text{s.t.} \quad & y = b - Ax \end{aligned}$$

$x \in \mathcal{D} \iff y \in \mathbb{R}^m_{++}$

The associated Lagrangian is (for $\nu \in \mathbb{R}^m$, and $x \in \mathcal{D}, y \in \mathbb{R}^m_{++}$)

$$\begin{aligned} L(x, y, \nu) &= -\sum_{i=1}^{m} \log(y_i) + \nu^T(y - b + Ax) \\ &= -\sum_{i=1}^{m} \log(y_i) + \nu^T y + \nu^T A x - \nu^T b \end{aligned}$$

Let $f : y \in \mathbb{R}^m_{++} \mapsto -\sum_{i=1}^m \log y_i + \nu^T y$. (We suppose $\nu_i > 0 \ \forall i = 1, \dots, m$)

On its domain, f is differentiable and one has:
$(\nabla f(y))_i = \nu_i - 1/y_i$, which is differentiable.
The hessian of f is $Hf(y) = \begin{pmatrix} 1/y_1^2 & & \\ & \ddots & \\ & & 1/y_m^2 \end{pmatrix} \succ 0$

Therefore f is convex and is such that $\nabla f(y_0) = 0$ for
$y_0 = (1/\nu_1, \dots, 1/\nu_m)^T$
So in the case $\nu > 0$, f reaches a minimum in y_0
that is $\min_{y \in \mathbb{R}^m_{++}} f(y) = m + \sum_{i=1}^m \log(\nu_i)$

If $\nu \not\succ 0$: $\exists i = 1, \dots, m$ s.t. $\nu_i \leq 0$
then let $y_t = (1, \dots, 1, \underbrace{t}_{i\text{th entry}}, 1, \dots, 1)$ $(t > 0)$

$f(y_t) = -\log(t) + \nu_i t + \text{cte}$
$\xrightarrow[t \to +\infty]{} -\infty$ so f is unbounded bellow.

Therefore, the dual function is:
$$g(\nu) = \begin{cases} m + \sum_{i=1}^m \log(\nu_i) - \nu^T b & \text{if } \begin{cases} \nu^T A = 0 \\ \nu \succ 0 \end{cases} \\ -\infty & \text{otherwise} \end{cases}$$

And the dual problem reads:
$$\begin{aligned} \max \quad & m + \sum_{i=1}^m \log(\nu_i) - \nu^T b \\ \text{s.t.} \quad & \nu^T A = 0 \\ & \nu \succ 0 \end{aligned}$$

Exercise 5

(1) : $\min \; f(x)$ s.t. $Ax = b$, $\phi(x) = f(x) + \alpha \|Ax - b\|_2^2$

($f : \mathbb{R}^n \to \mathbb{R}$ convex and differentiable , $A \in \mathbb{R}^{m \times n}$, $\mathrm{rank}(A) = m$)

We suppose $\tilde{x}$ minimizes ϕ. Since f is differentiable, so is ϕ.
Therefore $\nabla \phi(\tilde{x}) = 0 \Rightarrow \nabla f(\tilde{x}) + 2\alpha A^T(A\tilde{x} - b) = 0$

$$\Rightarrow \nabla f(\tilde{x}) + A^T \underbrace{(2\alpha(A\tilde{x} - b))}_{=\tilde{\nu}} = 0$$

The Lagrangian of (1) is $L(x, \nu) = f(x) + \nu^T(Ax - b)$
which is convex and differentiable.

$$\nabla_x L(x, \nu) = \nabla f(x) + A^T \nu$$

Hence, $\tilde{\nu} = 2\alpha(A\tilde{x} - b)$ is dual feasible and

$$g(\tilde{\nu}) = f(\tilde{x}) + \tilde{\nu}^T(A\tilde{x} - b)$$
$$= f(\tilde{x}) + 2\alpha \|A\tilde{x} - b\|_2^2 \leq f(x) \quad \forall x, \; Ax = b$$

$g(\tilde{\nu})$ is a lower bound on the optimal value of (1)

Exercise 6

See Jupyter Notebook