SELECTED SOLUTIONS, SECTION 3.2

1. Prove the Lagrangian sufficient conditions (3.2.3).

We consider the optimisation problem

$$\inf\{f(x) : g_i(x) \le 0 \text{ for } i = 1, 2, \dots, m\},$$

assume that $\bar{x}$ is feasible and that there exists a Lagrange multiplier λ. This implies that

$$f(\bar{x}) + \lambda^T g(\bar{x}) \le f(y) + \lambda^T y$$

for all $y \in \mathbb{E}$. Moreover, $\lambda_i = 0$ whenever $g_i(\bar{x}) < 0$ and $\lambda_i \ge 0$ and $g_i(\bar{x}) \le 0$ for all i. In particular, $\lambda^T g(\bar{x}) = 0$.

Now let $y \in \mathbb{E}$ be arbitrary but feasible. Then $\lambda^T g(y) \ge 0$. Thus

$$f(\bar{x}) = f(\bar{x}) + \lambda^T g(\bar{x}) \le f(y) + \lambda^T g(y) \le f(y).$$

Thus $\bar{x}$ is a solution of the constraint optimisation problem. (NB: convexity of either f or the constraint set has nowhere been used; even more: we have not even used the fact that $\mathbb{E}$ is a vector space!)

4. For a matrix A in $\mathbb{S}^n_{++}$ and a real $b > 0$, use the Lagrangian sufficient conditions to solve the problem

$$\inf\{-\log\det X : \operatorname{tr} AX \le b,\ X \in \mathbb{S}^n_{++}\}.$$

You may use the fact that the objective function is convex with derivative $-X^{-1}$ (see Section 3.1, Exercise 21 (The log barrier)).

We have to find $\lambda \ge 0$ such that $X \in \mathbb{S}^n_{++}$ satisfying $\operatorname{tr} AX \le b$ solves the optimisation problem

$$\inf\{-\log\det X + \lambda(\operatorname{tr} AX - b)\}.$$

Since the mapping $X \mapsto \operatorname{tr} AX - b$ is linear, the objective function $L(X, \lambda)$ of this problem is convex and differentiable with respect to X, and thus X is a minimiser if and only if $\nabla_X L(X, \lambda) = 0$. Now

$$\nabla_X L(X, \lambda) = -X^{-1} + \lambda A.$$

Thus we obtain the conditions that $X^{-1} = \lambda A$ and $\operatorname{tr} AX = b$ (since the assumption $\operatorname{tr} AX < b$ would imply that $\lambda = 0$ and thus $X^{-1} = 0$, which is impossible). From this we obtain that $X = bA^{-1}/n$ is a minimizer of $L(X, \lambda)$ with Lagrange multiplier $\lambda = n/b$.

6. (Extended convex function)
   (a) Give an example of a convex function that takes the values 0 and $-\infty$.
   (b) Prove the value function v defined by equation (3.2.4) is convex.
   (c) Prove that a function $h\colon \mathbb{E} \to [-\infty, \infty]$ is convex if and only if it satisfies the inequality

$$h(\lambda x + (1-\lambda)y) \le \lambda h(x) + (1-\lambda)h(y)$$

for any points x and y in $\operatorname{dom} h$ (or $\mathbb{E}$ if h is proper) and any real λ in $(0, 1)$.

Date: October 20, 2015.

(d) Prove that if the function $h\colon \mathbb{E} \to [-\infty, +\infty]$ is convex, then $\mathrm{dom}(h)$ is convex.

(a) Probably the simplest possible function is $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} +\infty & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -\infty & \text{if } x < 0. \end{cases}$$

The epigraph of this function is the set

$$\mathrm{epi}_f = \big\{(x,t) \in \mathbb{R}^2 : x > 0\big\} \cup \big\{(0,t) \in \mathbb{R}^2 : t \geq 0\big\},$$

which is a convex (but non-closed) set.

(b) Let (a,α), $(b,\beta) \in \mathrm{epi}_v$ and let $0 < \lambda < 1$. Then there exist for every $\varepsilon > 0$ some x_ε, $y_\varepsilon \in \mathbb{E}$ such that $g(x_\varepsilon) \leq a$, $g(y_\varepsilon) \leq b$, and $f(x_\varepsilon) \leq \alpha + \varepsilon$, $f(y_\varepsilon) \leq \beta + \varepsilon$. Now the convexity of g implies that

$$g(\lambda x_\varepsilon + (1-\lambda)y_\varepsilon) \leq \lambda g(x_\varepsilon) + (1-\lambda)g(y_\varepsilon) \leq \lambda a + (1-\lambda)b,$$

and the convexity of f implies that

$$g(\lambda x_\varepsilon + (1-\lambda)y_\varepsilon) \leq \lambda f(x_\varepsilon) + (1-\lambda)f(y_\varepsilon) \leq \lambda\alpha + (1-\lambda)\beta + \varepsilon.$$

Thus $\lambda x_\varepsilon + (1-\lambda)y_\varepsilon$ is feasible for $\inf\big\{f(x) : g(x) \leq \lambda a + (1-\lambda)b\big\}$, and thus

$$\begin{aligned} v(\lambda a + (1-\lambda)b) &= \inf\big\{f(x) : g(x) \leq \lambda a + (1-\lambda)b\big\} \\ &\leq \inf_{\varepsilon>0}(\lambda f(x_\varepsilon) + (1-\lambda)f(y_\varepsilon)) \\ &= \lambda\alpha + (1-\lambda)\beta. \end{aligned}$$

This shows that $\lambda(a,\alpha) + (1-\lambda)(b,\beta) \in \mathrm{epi}_v$, and thus v is convex.

(c) Assume first that $h\colon \mathbb{E} \to [-\infty,\infty]$ is convex and let x, $y \in \mathrm{dom}(h)$ and $0 < \lambda < 1$. Let moreover ξ, $\eta \in \mathbb{R}$ satisfy $\xi \geq h(x)$ and $\eta \geq h(y)$. Then (x,ξ), $(y,\eta) \in \mathrm{epi}_h$ and thus the convexity of epi_h implies that $\lambda(x,\xi) + (1-\lambda)(y,\eta) \in \mathrm{epi}_h$, which means that

$$h(\lambda x + (1-\lambda)y) \leq \lambda\xi + (1-\lambda)\eta.$$

As a consequence,

$$h(\lambda x+(1-\lambda)y) \leq \inf\big\{\lambda\xi+(1-\lambda)\eta : \xi \geq h(x) \text{ and } \eta \geq h(y)\big\} = \lambda h(x)+(1-\lambda)h(y).$$

(Note that we cannot choose $\xi = h(x)$ and $\eta = h(y)$, because $h(x)$ and $h(y)$ might be $-\infty$!)

Conversely, assume that the inequality given in this exercise holds and let (x,ξ), $(y,\eta) \in \mathrm{epi}_h$. Since $h(x) \leq \xi < \infty$ and $h(y) \leq \eta < \infty$, it follows that x, $y \in \mathrm{dom}(h)$. Now, if $0 < \lambda < 1$ we have

$$h(\lambda x + (1-\lambda)y) \leq \lambda h(x) + (1-\lambda)h(y) \leq \lambda\xi + (1-\lambda)\eta,$$

showing that $\lambda(x,\xi) + (1-\lambda)(y,\eta) \in \mathrm{epi}_h$. Thus h is convex.

(d) Let x, $y \in \mathrm{dom}(h)$ and let $0 < \lambda < 1$. Then (by (c))

$$h(\lambda x + (1-\lambda)y) \leq \lambda h(x) + (1-\lambda)h(y) < \infty,$$

and thus $\lambda x + (1-\lambda)y \in \mathrm{dom}(h)$.

7. (Nonexistence of multiplier) For the function $f\colon \mathbb{R} \to (-\infty, +\infty]$ defined by $f(x) = -\sqrt{x}$ for $x \in \mathbb{R}_+$ and $+\infty$ otherwise, show there is no Lagrange multiplier at the optimal solution of $\inf\big\{f(x) : x \leq 0\big\}$.

First note that the solution of $\inf\{f(x) : x \le 0\}$ is $x = 0$ (this is the only feasible point in the domain of f). Now consider the Lagrangian

$$L(x, \lambda) = -\sqrt{x} + \lambda x$$

for $x \ge 0$ and $\lambda \ge 0$. First we note that $\inf_{x\ge 0} L(x, 0) = -\infty$, and thus $\lambda = 0$ is no Lagrange multiplier at x. On the other hand, if $\lambda > 0$, then

$$L(\lambda^{-2}/4, \lambda) = -\lambda/2 + \lambda/4 = -\lambda/4 < 0 = L(0, \lambda),$$

and thus $x = 0$ is no global minimum of $L(\cdot, \lambda)$.

11. (Normals to epigraphs) For a function $f : \mathbb{E} \to (-\infty, \infty]$ and a point $\bar{x}$ in $\operatorname{core}(\operatorname{dom} f)$, calculate the normal cone $N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$.

We recall that the normal cone $N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$ consists of all $(d, \tau) \in \mathbb{E} \times \mathbb{R}$ such that

$$\langle d, y - \bar{x}\rangle + \tau(\eta - f(\bar{x})) \le 0$$

whenever $y \in \operatorname{dom}(f)$ and $\eta \ge f(y)$.

We now discuss the different possibilities for τ separately:

- Choosing $y = \bar{x}$ and $\eta \ge f(\bar{x})$ arbitrary, this implies that, necessarily, $\tau \le 0$.
- Now assume that $\tau = 0$ and $(d, 0) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$. Then

  $$\langle d, y - \bar{x}\rangle \le 0$$

  for every $y \in \operatorname{dom}(f)$. However, since $\bar{x} \in \operatorname{core}(\operatorname{dom} f)$, we may choose $y = \bar{x} - td$ for some $t > 0$. Thus $0 \ge \langle d, td\rangle = t\|d\|^2$ and thus $d = 0$.
- Finally assume that $\tau < 0$ and $(d, \tau) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$. Then we have in particular that

  $$\langle d, y - \bar{x}\rangle \le -\tau(f(y) - f(\bar{x}))$$

  whenever $y \in \operatorname{dom}(f)$. Dividing by $-\tau > 0$, it follows that

  $$\langle -d/\tau, y - \bar{x}\rangle \le f(y) - f(\bar{x})$$

  for every $y \in \operatorname{dom}(f)$. Obviously, the same inequality also holds for $y \notin \operatorname{dom}(f)$, because then the right hand side is infinite. This, however, implies that $-d/\tau$ is a subgradient of f at the point $\bar{x}$.

Thus $N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$ consists of all vectors (d, τ) such that either $d = 0$ and $\tau = 0$, or $\tau < 0$ and $-d/\tau \in \partial f(\bar{x})$. Put differently,

$$N_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) = \{(0, 0)\} \cup \mathbb{R}_+(\partial f(\bar{x}), -1).$$

Note: The argumentation above applies also, if the function f is non-convex, provided that one uses the following definitions:

- If $C \subset \mathbb{E}$ and $x \in C$, one defines the normal cone $N_C(x)$ to C at x as the set of all directions $d \in \mathbb{E}$ such that

  $$\langle x, x - \bar{x}\rangle \le 0 \qquad \text{for all } x \in C.$$

- If $f : \mathbb{E} \to (-\infty, +\infty]$ is a function and $x \in \operatorname{dom}(f)$, then the subgradient of f at x consists of all $\phi \in \mathbb{E}$ such that

  $$\langle \phi, x - \bar{x}\rangle \le f(x) - f(\bar{x}) \qquad \text{for all } x \in \mathbb{E}.$$