ESI 6420 HOMEWORK 4 SOLUTIONS

MAKSIM LEVENTAL

Time spent: 15 hours

Collaborators: Chris Gianelli

2.1 Claim: for a continuous $f : \mathbb{R} \to \mathbb{R}$ and $h > 0$

$$f_h(x) := \frac{1}{2h} \int_{x-h}^{x+h} f(t)\, dt \geq f(x)$$

iff f is convex.

Proof. $\Leftarrow$ Suppose f is convex. Proceed by contradiction: suppose there exist h_0, x_0 such that $f(x_0) > f_{h_0}(x_0)$. Since f is convex there exists $g(x) = f(x_0) + m(x - x_0)$ such that $g \leq f$. But then

$$f(x_0) = \frac{1}{2h} \int_{x_0-h_0}^{x_0+h_0} g(t)\, dt \leq \frac{1}{2h} \int_{x_0-h_0}^{x_0+h_0} f(t)\, dt = f_{h_0}(x_0)$$

a contradiction.

$\Rightarrow$ Suppose $f_h(x) \geq f(x)$ for all h, x. Towards a contradiction suppose f is not convex. Then there exist λ_0, x_1, x_2 such that

$$f(\lambda_0 x_1 + (1 - \lambda_0) x_2) > \lambda_0 f(x_1) + (1 - \lambda_0) f(x_2)$$

where $\lambda_0 \in (0, 1)$. Consider the function

$$F(\lambda) = f(\lambda x_1 + (1 - \lambda) x_2) - (\lambda f(x_1) + (1 - \lambda) f(x_2))$$

on $[0, 1]$. Note that $F(0) = F(1) = 0$ and $F(\lambda_0) > 0$. Therefore there exists an h-ball around λ_0 such that for $\lambda \in [\lambda_0 - h, \lambda_0 + h]$ it's the case that $F(\lambda) > 0$. Note also that F is continuous because it's a linear function of a continuous function f. Therefore F achieves a maximum for some$\lambda^* \in [\lambda_0 - h, \lambda_0 + h]$ (since $[\lambda_0 - h, \lambda_0 + h]$ is closed). Take a smaller ball $[\lambda^* - h', \lambda^* + h']$ around λ^* and suppose F is not constant on this smaller ball (we'll relax this in a moment). Then since F is positive on $[\lambda^* - h', \lambda^* + h']$ we have that

$$2h' F(\lambda^*) > \int_{\lambda^*-h'}^{\lambda^*+h'} F(\lambda)\, d\lambda$$

which is equivalent to

$$\begin{aligned} & f(\lambda^* x_1 + (1 - \lambda^*) x_2) - \\ & (\lambda^* f(x_1) + (1 - \lambda^*) f(x_2)) > \frac{1}{2h'} \int_{\lambda^*-h'}^{\lambda^*+h'} [f(\lambda x_1 + (1 - \lambda) x_2) - (\lambda f(x_1) + (1 - \lambda) f(x_2))]\, d\lambda \\ & = \frac{1}{2h'} \int_{\lambda^*-h'}^{\lambda^*+h'} f(\lambda x_1 + (1 - \lambda) x_2)\, d\lambda - (\lambda^* f(x_1) + (1 - \lambda^*) f(x_2)) \end{aligned}$$

Cancelling $-(\lambda^* f(x_1) + (1 - \lambda^*) f(x_2))$ we get that

$$f(\lambda^* x_1 + (1 - \lambda^*) x_2) > \frac{1}{2h'} \int_{\lambda^*-h'}^{\lambda^*+h'} f(\lambda x_1 + (1 - \lambda) x_2)\, d\lambda$$

$\square$

Date: November 20, 2015.

2.2 Let $f(X) = -\log(\det(X))$.

(a) Claim: For $X, D \succeq 0$ and $X \succ 0$ and $g(t) = f(X + tD)$ it's the case that

$$g(t) = -\log\left(\det\left(\sqrt{X}\left(I + t\left(\sqrt{X}\right)^{-1} D\left(\sqrt{X}\right)^{-1}\right)\sqrt{X}\right)\right)$$

Proof. Firstly since $X \succ 0$ it's the case that X is full rank (all nonzero eigenvalues) and there exists a matrix $\sqrt{X}$ such that $\sqrt{X}\sqrt{X} = X$ and $\sqrt{X}$ is full rank ($\sqrt{X} = Q\sqrt{\Sigma}Q^T$ where Q is the set of eigenvectors corresponding to X and $\sqrt{\Sigma} \succ 0$ since $\Sigma \succ 0$). Then $\left(\sqrt{X}\right)^{-1}$ exists and hence

$$\sqrt{X}\left((I + t)\left(\sqrt{X}\right)^{-1} D\left(\sqrt{X}\right)^{-1}\right)\sqrt{X} = X + tD$$

and so $g(t) = -\log(\det(X + tD)) = f(X + tD)$. □

(b) Claim: $f(X)$ is convex.

Proof. Using the representation of g proven to be appropriate in part (a)

$$\begin{aligned} g(t) &= -\log\left(\det\left(\sqrt{X}\right)\det\left(I + t\left(\sqrt{X}\right)^{-1} D\left(\sqrt{X}\right)^{-1}\right)\det\left(\sqrt{X}\right)\right) \\ &= -\log(\det(X)) - \log\left(\det\left(I + t\left(\sqrt{X}\right)^{-1} D\left(\sqrt{X}\right)^{-1}\right)\right) \end{aligned}$$

Let $Y = \left(\sqrt{X}\right)^{-1} D\left(\sqrt{X}\right)^{-1}$, which is PSD since D is PSD and $\left(\sqrt{X}\right)^{-1}$ is PD, and

$$\begin{aligned} g(t) &= -\log(\det(X)) - \log(\det(I + tY)) \\ &= -\log(\det(X)) - \log\left(\prod_{i=1}^{n}(1 + t\lambda_i)\right) \\ &= -\log(\det(X)) - \sum_{i=1}^{n}\log(1 + t\lambda_i) \end{aligned}$$

Note g is convex in t since it's the sum of a constant and convex functions of linear transformations of t. Therefore $f(X)$ is convex since it is convex on every line. □

(1) Claim: Let $c \sim \mathcal{N}(\mu, \Sigma)$. Then assuming there exists x such that $P(c^\intercal x \geq \alpha) \geq \frac{1}{2}$

$$\begin{aligned} &\max_{x \in \mathbb{R}^n} P(c^\intercal x \geq \alpha) \\ &\text{s.t.} Fx \leq g \\ &\quad Ax = b \end{aligned}$$

can be reformulated as a quadratic convex optimization problem.

Proof. Firstly since $c \sim \mathcal{N}(\mu, \Sigma)$ it's the case that $X = c^\intercal x \sim \mathcal{N}(\mu \cdot x, x^\intercal \Sigma x)$ and hence

$$P(X \geq \alpha) = P\left(\frac{X - \mu \cdot x}{\sqrt{x^\intercal \Sigma x}} \geq \frac{\alpha - \mu \cdot x}{\sqrt{x^\intercal \Sigma x}}\right) = P\left(Z \geq \frac{\alpha - \mu \cdot x}{\sqrt{x^\intercal \Sigma x}}\right)$$

where $Z \sim \mathcal{N}(0, 1)$. So the maximization problem is

$$\max_x \left[P\left(Z \geq \frac{\alpha - \mu \cdot x}{\sqrt{x^\intercal \Sigma x}}\right)\right]$$

Clearly maximizing this objective is equivalent to minimizing $\frac{\alpha - \mu \cdot x}{\sqrt{x^\intercal \Sigma x}}$. The fact that there exists x_0 such that

$$P(c^\intercal x_0 \geq \alpha) \geq \frac{1}{2}$$

means there exists x_0 such that

$$P\left(Z \geq \frac{\alpha - \mu \cdot x_0}{\sqrt{x_0^\intercal \Sigma x_0}}\right) \geq \frac{1}{2}$$

and so there exists x_0 such that

$$\frac{\alpha - \mu \cdot x_0}{\sqrt{x_0^\intercal \Sigma x_0}} \leq 0$$

or

$$\alpha - \mu \cdot x_0 \leq 0$$

So the problem now is

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} \frac{\alpha - \mu \cdot x}{\sqrt{x^\intercal \Sigma x}} \\ &\text{s.t.} Fx \leq g \\ &Ax = b \end{aligned}$$

Let $z = \frac{1}{\sqrt{x^\intercal \Sigma x}}$ and $y = \frac{x}{\sqrt{x^\intercal \Sigma x}}$. Then the minimization problem is

$$\begin{aligned} &\min_{(z,y) \in \mathbb{R}^n} \alpha z - \mu \cdot y \\ &\text{s.t.} Fx \leq g \\ &Ax = b \end{aligned}$$

$\square$