MATH 680: Assignment 2

Annik Gougeon, David Fleischer

Last Update: 21 February, 2018

Question 1

Proof. Let $C = \{x \in \mathbb{R}^n : Ax \leq b\}$ be our set of interest. Let $x, y \in C$, and let $t \in [0,1]$ be an arbitrary real-valued scalar. Then,

$$\begin{aligned} A\left(tx + (1-t)y\right) &= tAx + (1-t)Ay \\ &\leq tb + (1-tb) \\ &= b. \end{aligned}$$

Thus,

$$x, y \in C \implies tx + (1-t)y \in C, \quad \text{for all } 0 \leq t \leq 1.$$

That is, C is a convex set, as desired. $\square$

Question 2

2.1

Proof. Recall that a (continuous, twice differentiable) function $f(z)$, $z \in C$, is convex on C if and only if its Hessian is positive semidefinite for all z on the interior of C,

$$\nabla^2 f(z) \in \mathbb{S}^n_+,$$

and strongly convex with parameter $m > 0$ if and only if

$$\nabla^2 f(z) - m\mathbb{I}_n \in \mathbb{S}^n_+.$$

Furthermore, a matrix $M \in \mathbb{R}^{n\times n}$ is positive semidefinite if and only if all eigenvalues of M are nonnegative. Since f is nondifferentiable along $x = 0, y = 0$ we first apply a differentiable approximation f_ϵ

$$f_\epsilon(x,y) = \sqrt{x^2y^2 + \epsilon} + a\left(x^2 + y^2\right) \underset{\epsilon \to 0}{\longrightarrow} |xy| + a\left(x^2 + y^2\right) = f(x,y)$$

Now, f_ϵ admits gradient

$$\nabla f_\epsilon(x,y) = \left(2ax + \frac{xy^2}{\sqrt{x^2y^2+\epsilon}}, 2ay + \frac{x^2y}{\sqrt{x^2y^2+\epsilon}}\right),$$

and Hessian

$$\nabla^2 f_\epsilon(x,y) = \begin{bmatrix} -\frac{x^2y^4}{(x^2y^2+\epsilon)^{3/2}} + \frac{y^2}{\sqrt{x^2y^2+\epsilon}} + 2a & \frac{2xy}{\sqrt{x^2y^2+\epsilon}} - \frac{x^3y^3}{(x^2y^2+\epsilon)^{3/2}} \\ \frac{2xy}{\sqrt{x^2y^2+\epsilon}} - \frac{x^3y^3}{(x^2y^2+\epsilon)^{3/2}} & -\frac{y^2x^4}{(x^2y^2+\epsilon)^{3/2}} + \frac{x^2}{\sqrt{x^2y^2+\epsilon}} + 2a \end{bmatrix}.$$

We find $\nabla^2 f_\epsilon(x,y)$ to be[1]

$$\lambda_{\epsilon,1} = \frac{x^2\left(4ay^2\sqrt{x^2y^2+\epsilon}+\epsilon\right)+4a\epsilon\sqrt{x^2y^2+\epsilon}-\sqrt{4x^6y^6+x^4\epsilon\left(16y^4+\epsilon\right)+14x^2y^2\epsilon^2+y^4\epsilon^2}+y^2\epsilon}{2\left(x^2y^2+\epsilon\right)^{3/2}}$$

$$\lambda_{\epsilon,2} = \frac{x^2\left(4ay^2\sqrt{x^2y^2+\epsilon}+\epsilon\right)+4a\epsilon\sqrt{x^2y^2+\epsilon}+\sqrt{4x^6y^6+x^4\epsilon\left(16y^4+\epsilon\right)+14x^2y^2\epsilon^2+y^4\epsilon^2}+y^2\epsilon}{2\left(x^2y^2+\epsilon\right)^{3/2}}.$$

Taking the limits of $\lambda_{\epsilon,1}$ and $\lambda_{\epsilon,2}$ as $\epsilon \to 0$,

$$\begin{aligned}
\lambda_1 = \lim_{\epsilon\to 0}\lambda_{\epsilon,1} &= \frac{4ax^2y^2\sqrt{x^2y^2}-2\sqrt{x^6y^6}}{2\left(x^2y^2\right)^{3/2}} \\
&= 2a - \frac{\left(x^2y^2\right)^{3/2}}{\sqrt{x^6y^6}} \\
&= 2a-1, \\
\lambda_2 = \lim_{\epsilon\to 0}\lambda_{\epsilon,2} &= \frac{4ax^2y^2\sqrt{x^2y^2}+2\sqrt{x^6y^6}}{2\left(x^2y^2\right)^{3/2}} \\
&= 2a + \frac{\left(x^2y^2\right)^{3/2}}{\sqrt{x^6y^6}} \\
&= 2a+1.
\end{aligned}$$

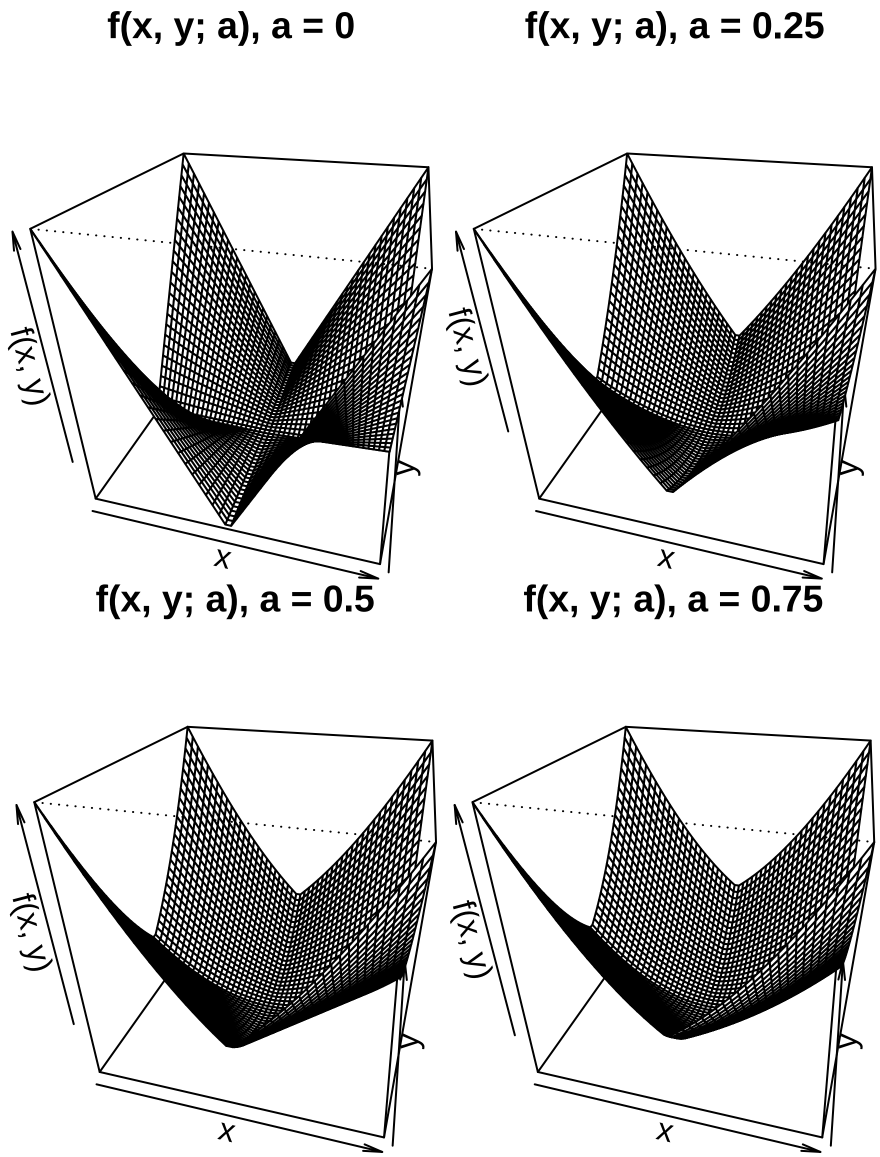
In this form we see that $\nabla^2 f(x,y)$ has nonnegative eigenvalues if and only if $a \geq \frac{1}{2}$, and so f is convex for $a \geq \frac{1}{2}$. To show strong convexity, we use the result that if matrix M has eigenvalues $\{\lambda_i\}_{i=1}^n$ then $M - k\mathbb{I}_n$ has eigenvalues $\{\lambda_i - k\}_{i=1}^n$. Therefore, $\nabla^2 f(x,y) - m\mathbb{I}_2$ has eigenvalues

$$\begin{aligned}
\lambda_{m,1} &= 2a-1-m \\
\lambda_{m,2} &= 2a+1-m.
\end{aligned}$$

To ensure $\lambda_{m,1}, \lambda_{m,2}$ are nonnegative we set $a > \frac{1}{2}$ and $m \leq a$. Therefore, f is strongly convex with parameter m, $a \geq m > 0$, as desired. $\square$

We present below figures of f evaluated on $[-1,1]\times[-1,1]$ for $a \in \{0, 0.25, 0.5, 0.75\}$.

[1] Proof left as an exercise.



2.2

2.2 (a)

For $x \in \mathbb{R}^n_{++}$ we find gradient

$$\nabla f(x) = -\left[x_1^{-1}, x_2^{-1}, \cdots, x_n^{-1}\right],$$

and Hessian

$$\nabla^2 f(x) = \begin{bmatrix} x_1^{-2} & 0 & \cdots & 0 \\ 0 & x_2^{-2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n^{-2} \end{bmatrix}.$$

Since $\nabla^2 f(x)$ is diagonal we may immediately obtain its eigenvalues $\{\lambda_i\}_{i=1}^n$,

$$\lambda_i = x_i^{-2}.$$

We see that, since $x \in \mathbb{R}^n_{++} \iff x_i > 0, i = 1, ..., n$, all eigenvalues $\lambda_i > 0$. Therefore, f must be strongly convex (and so strictly convex, and convex), as desired.

2.2 (b)

2.3

Proof. ($\implies$) Suppose f is convex. Then, $\mathrm{dom}(f)$ is a convex set, and, for all $x, y \in \mathrm{dom}(f)$ and $t \in [0, 1]$,

$$\begin{aligned} f(tx + (1-t)y) &\le tf(x) + (1-t)f(y) \\ \iff f(t(x-y) + y) - f(y) &\le t\,(f(x) - f(y)) \\ \iff \frac{f(t(x-y)+y) - f(y)}{t} &\le f(x) - f(y) \\ \iff \frac{f(t(x-y)+y) - f(y)}{t} + f(y) &\le f(x) \end{aligned}$$

Note that if we take the limit of our first term as $t \to 0$, for finite x, y,

$$\begin{aligned} \lim_{t\to 0} \frac{f(t(x-y)+y) - f(y)}{t} &= \frac{\partial}{\partial t} f(t(x-y)+y)\Big]_{t=0} \\ &= \nabla f(t(x-y)+y)^T(x-y)\Big]_{t=0} \\ &= \nabla f(y)^T(x-y). \end{aligned}$$

Therefore, taking the limit of our inequality above as $t \to 0$,

$$\begin{aligned} \lim_{t\to 0}\left(\frac{f(t(x-y)+y) - f(y)}{t} + f(y)\right) &\le \lim_{t\to 0} f(x) \\ \iff \nabla f(y)^T(x-y) + f(y) &\le f(x). \end{aligned}$$

By symmetry we swap x and y to obtain

$$f \text{ convex} \implies \mathrm{dom}(f) \text{ convex and } f(y) \ge f(x) + \nabla f(x)^T(y-x),$$

as desired.

($\Longleftarrow$) Suppose $\mathrm{dom}(f)$ is convex and, for $x, y \in \mathrm{dom}(f)$, $x \ne y$,

$$\nabla f(x)^T(y-x) + f(x) \le f(y).$$

Since $\mathrm{dom}(f)$ is convex we find $z = tx + (1-t)y \in \mathrm{dom}(f)$, $t \in [0, 1]$. Then, for such x, y, z,

$$\begin{aligned} \nabla f(z)^T(x-z) + f(z) &\le f(x) \\ \nabla f(z)^T(y-z) + f(z) &\le f(y). \end{aligned}$$

Multiplying our first inequality by t and the second by $(1-t)$, and then adding the two yields

$$\begin{aligned} t\left[\nabla f(z)^T(x-z) + f(z)\right] + (1-t)\left[\nabla f(z)^T(y-z) + f(z)\right] &\le tf(x) + (1-t)f(y) \\ \iff t\nabla f(z)^T(x-z) + (1-t)\nabla f(z)^T(y-z) + f(z) &\le tf(x) + (1-t)f(y) \\ \iff \nabla f(z)^T\left[t(x-z) + (1-t)(y-z)\right] + f(z) &\le tf(x) + (1-t)f(y) \\ \iff \nabla f(z)^T\left[tx + (1-t)y - z\right] + f(z) &\le tf(x) + (1-t)f(y) \\ \iff f(tx + (1-t)y) &\le tf(x) + (1-t)f(y), \end{aligned}$$

where the final line was achieved by recalling that $z = tx + (1-t)y$. Therefore,

$$f \text{ convex} \Longleftarrow \text{dom}(f) \text{ convex and } f(y) \geq f(x) + \nabla f(x)^T(y-x),$$

as desired. □

Question 3

(a)

Proof. (1 $\implies$ 2) (2 $\implies$ 3) (3 $\implies$ 4) (4 $\implies$ 1)

□

(b)

Proof. Let f be convex and twice differentiable.

(1 $\implies$ 2) If ∇f is L-Lipschitz then, for $x, y \in \text{dom}(f)$, $L > 0$,

$$\|\nabla^2 f(x) - \nabla^2 f(y)\|_2 \leq L\|x-y\|_2.$$

(2 $\implies$ 3) (3 $\implies$ 4) (4 $\implies$ 1)

□

Question 4