CS 726 Assignment 1

Ruochen Lin

February 9, 2018

1

1.1

This is a polyhedron with

$$Ix \geq 0$$
$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_n \\ a_1^2 & a_2^2 & \dots & a_n^2 \end{bmatrix} x = \begin{bmatrix} 1 \\ b \\ c \end{bmatrix}.$$

1.2

$x^T y \leq \|x\|_2 \|y\|_2$; the inequality becomes equality when $y = \alpha x$, $\alpha > 0$. Thus from $x^T y \leq 1$, $\forall y$ s.t. $\|y\|_2 = 1$ we can know $\|x\|_2 \leq 1$. This is a hypersphere in $\mathbb{R}^n$, which is not a polyhedron.

1.3

For all $\{y \,|\, \|y\|_1 = 1\}$, $x^T y \leq \|x\|_\infty$; the inequality becomes equality when y has entry 1 at the position corresponding to $\max\{x_i\}$ and zeros else where. Thus the condition given is equivalent to $\|x\|_\infty \leq 1$. The resulting shape is a polyhedron, given by the conditions of $\{x \,|\, -1 \leq x_i \leq 1\}$; this can be represented by

$$Ix \geq \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix}$$

and

$$-Ix \geq \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix}.$$

If we further take $x \geq 0$ into consideration, we have

$$Ix \geq 0$$

in place of the first inequality. In addition, $F = 0$ and $g = 0$.

2

If x^* is a local minimum, then there $\exists \mathcal{N}(x^*) : \forall x \in \mathcal{N}(x^*), f(x) \geq x^*$. In addition, if x^* is not a strict local minimum, $\forall \mathcal{N}(x^*), \exists x^\dagger \in \mathcal{N}(x^*) : f(x^\dagger) \leq f(x^*)$. The two conditions combines to state that, in those $\mathcal{N}(x^*)$ that make x^* a local minimum, $x^\dagger = x^*$. Also, $x^\dagger$ must not be on the boundaries of those $\mathcal{N}(x^*)$; otherwise if we leave out $x^\dagger$ in $\mathcal{N}(x^*)$, it will be a neighbourhood that makes x^* a strict minimum. Thus in (x^*), which is now also $\mathcal{N}(x^\dagger)$, $\forall x \in \mathcal{N}(x^\dagger)$, $f(x) \geq f(x^\dagger) = f(x^*)$. This means that $x^\dagger$ is also a local minimum, and $x^\dagger$ exists in all neighbourhoods of x^*; hence we prove that x^* cannot be an isolated minimum if it is not a strict one. Its contraposition, which states that all isilated minima are strict, is also true.

3

3.1

A square matrix that has all of its entries being 1 is not positive definite because 0 is one of its eigenvalues, despite having all positive entries.

3.2

Yes. If one of its diagonal entries is nonpositive, say $a_{ii} \leq 0$, then $x^T Ax = a_{ii}x_i^2 \leq 0$ with $x_j = \delta_{ij}$. Note that δ_{ij} is the Kronecker delta, which gives 1 when $i = j$ and 0 otherwise.

4

$$\nabla f(x) = \begin{bmatrix} 400x_1^3 - 400x_1x_2 + 2x_1 - 2 & -200x_1 + 200x_2 \end{bmatrix}^T$$

$$\nabla^2 f(x) = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

Inserting $x = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ into the expression of the gradient, we have $\nabla f(\begin{bmatrix} 1 & 1 \end{bmatrix}^T) = 0$, thus this is a stationary point of $f(x)$. In addition, at this point the Hessian matrix is

$$\nabla f(\begin{bmatrix} 1 & 1 \end{bmatrix}^T) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}.$$

The determinant of $\nabla^2 f(x)$ is 4, and the determinant of [802] is 802, both being positive. This shows that the Hesssian at the point is positive definite and the point is a local minimum.

5

$$\nabla f(x) = \begin{bmatrix} 2x_1 + \beta x_2 + 1 & 2x_2 + \beta x_1 + 2 \end{bmatrix}^T$$

$$\nabla^2 f(x) = \begin{bmatrix} 2 & \beta \\ \beta & 2 \end{bmatrix}$$

The solution to $\nabla f(x) = 0$ is $x = \begin{bmatrix} \frac{2\beta-2}{4-\beta^2} & \frac{\beta-4}{4-\beta^2} \end{bmatrix}^T$, given that $\beta \neq \pm 2$. To further make the solution global minimizer, we require the Hessian to be positively definite, so that the function is strictly convex. The eigenvalues of the Hessian are $\lambda = 2 \pm \beta$, and only when both eigenvalues are positive is the Hessian positive definite. To meet these criteria, we have $\lambda \in (-2, 2)$.

6

Define $p \equiv y - x$, then the characterization of convexity can be written as

$$f(x+p) \geq f(x) + \nabla f(x)^T p.$$

Use Taylor's theorem:

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x + \gamma p)p, \gamma \in [0, 1].$$

First, if the Hessian is positive semi-definite for $\forall x \in \mathbb{R}^n$, then $\frac{1}{2}p^T\nabla^2 f(x+\gamma p)p \geq 0$, $\forall p \in \mathbb{R}^n$. The remaining part of the equation is exactly the characterization inequality at its equality, which is now satisfied, hence proves f is convex if $\nabla^2 f(x) \succeq 0$. On the other hand, if $\nabla^2 f(x) \prec 0$ in some region of $\mathbb{R}^n$, then we can easily x in the region and p small enough so that both $y = x+p$ and all points on the intersect connecting the two points also falls in the region and that $\frac{1}{2}p^T\nabla^2 f(x+\gamma p)p < 0$. Then the characterization of convexity is violated, and thus f is not convex in the region.

7

7.1

$$\nabla f(x) = \begin{bmatrix} 4x_1^3 - 16x_1 & 2x_2 \end{bmatrix}^T$$

$$\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 - 16 & 0 \\ 0 & 2 \end{bmatrix}$$

$\nabla f(x) = 0$ has three solutions: $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$, $\begin{bmatrix} 2 & 0 \end{bmatrix}^T$, and $\begin{bmatrix} -2 & 0 \end{bmatrix}^T$. At $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$ the Hessian is:

$$\begin{bmatrix} -16 & 0 \\ 0 & 2 \end{bmatrix}^T,$$

which has a negative eigenvalue of -16, making it a first-order saddle point. At points $\begin{bmatrix} \pm 2 & 0 \end{bmatrix}^T$, the Hessian is

$$\begin{bmatrix} 32 & 0 \\ 0 & 2 \end{bmatrix}^T,$$

which is positive definite. This means that the latter two solutions are local minima of the function $f(x)$. Note that $f(x) \geq 0$ and $f(\begin{bmatrix} \pm 2 & 0 \end{bmatrix}^T) = 0$; hence the two local minima are global minima as well.

7.2

$$\nabla f(x) = \begin{bmatrix} x_1 + \cos x_2 & -x_1 \sin x_2 \end{bmatrix}^T$$

$$\nabla^2 f(x) = \begin{bmatrix} 1 & -\sin x_2 \\ -\sin x_2 & -x_1 \cos x_2 \end{bmatrix}$$

The solutions to $\nabla f(x) = 0$ are $\begin{bmatrix} 0 & (k+\frac{1}{2})\pi \end{bmatrix}^T$ and $\begin{bmatrix} (-1)^{k+1} & k\pi \end{bmatrix}^T$ ($k \in \mathbb{Z}$). For the first family of solutions,

$$\nabla^2 f(\begin{bmatrix} 0 & (k+\frac{1}{2})\pi \end{bmatrix}^T) = \begin{bmatrix} 1 & (-1)^{k+1} \\ (-1)^{k+1} & 0 \end{bmatrix} \not\succeq 0.$$

This matrix has eigenvalues $\lambda = \frac{1 \pm \sqrt{5}}{2}$, one of which is negative. Because the Hessian is not positive definite, these points are only saddle points of $f(x)$.
As for the latter family of solutions,

$$\nabla^2 f(\begin{bmatrix} (-1)^{k+1} & k\pi \end{bmatrix}^T) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \succ 0,$$

These are local minima of $f(x)$, and at these points $f(x) = \frac{1}{2}(-1)^{2(k+1)} + (-1)^{k+1} \cos k\pi = -\frac{1}{2}$.

7.3

$$\nabla f(x) = \begin{bmatrix} 4x_1^3 - 4x_1x_2 - 2x_1 & -2x_1^2 + 2x_2 \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} 12x_1 - 4x_2 - 2 & -4x_1 \\ -4x_1 & 2 \end{bmatrix}$$

There is only one solution to the equation $f(x) = 0$: $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$. At this point the Hessian is

$$\nabla^2 f(0) = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \not\succeq 0,$$

which has one negative eigenvalue and hence is not positive definite. Thus $f(x)$ has only one stationary point $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$, which is a saddle point.