UNIVERSITY OF CALIFORNIA, BERKELEY

HW5: EE127

Oscar Ortega

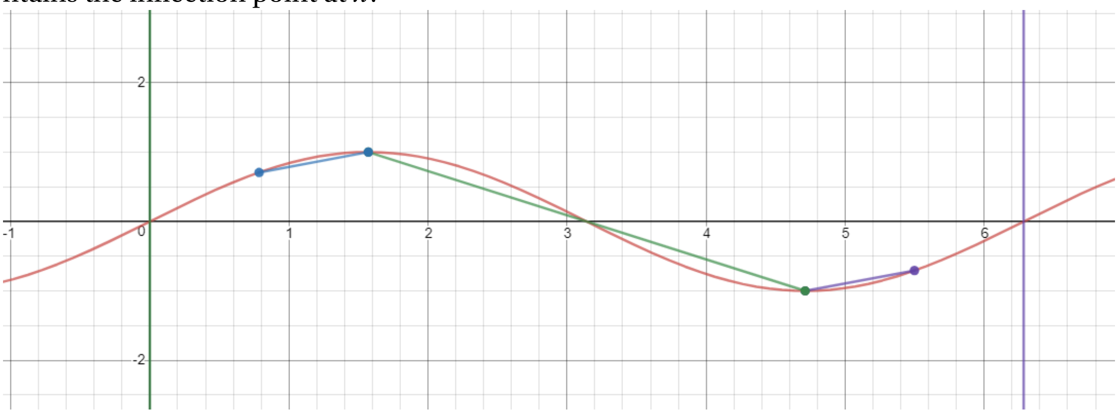
March 8, 2019

1 PROPERTIES OF CONVEX FUNCTION

a:

$$f(x) = sin(x)$$
$$f'(x) = cos(x)$$
$$f''(x) = -sin(x)$$

By the second-order conditions for convexity, we know that $f(x)$ is convex from the interval from $[\pi, 2\pi]$ because the second derivative is positive in this region. Similarly, the function is concave in $[0, \pi]$ because the second derivative of this function is negative. If we consider the region between $[\frac{\pi}{2}, \frac{3\pi}{2}]$ we can notice the function is neither convex nor concave because it contains the inflection point at π.



b:

We can notice that the graph of the chords will either remain above the function if the function is convex in the given interval, or strictly below if the function is concave. We can also

note that the chord will intersect the function if it lies in a region that is neither concave nor convex.
c: Consider $x \in [0, \pi/2]$

$$f_1(x) = \frac{2}{\pi}x$$

$$f_2(x) = sin(x)$$

$$f_3(x) = x$$

$$f_1(0) = f_2(0) \text{ and } f_1(\pi/2) = f_2(\pi/2)$$

which implies the function first function in the region $[0, \pi/2]$ is a chord of the concave function f_2. Therefore, it will be strictly less than or equal to second function $sin(x)$. We also know that if a function is concave and differentiable, the tangent line of any point of the function will be strictly greater than or equal to the concave function. Because we know that $sin(x)$ is concave on the interval $[0, 2\pi]$ we know that at $f_2(0)$ where $f_2'(0) = 1$ implies that $f_3(x)$ is the tangent line of $f_2(x)$ at $x = 0$, which implies that $f_3(x)$ is greater than $f_2(x)$ on the interval specified.

2 Epigraph and Convex Optimization

a: Proof:
let epi$f$ convex:
$(x_1, f(x_1)), (x_2, f(x_2) \in \text{epi}f$
$\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) \in \text{epi}f$
$\lambda f(x_1) + (1-\lambda)f(x_2) \geq f(\lambda(x_1) + (1-\lambda)x_2)$
$\rightarrow f$ is convex.
let f convex:
$f(\lambda(x_1) + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$
$\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) \in \text{epi}f$

epi$f$ is convex.
b: Proof: Let x^* be a local minimum and let $z = \lambda x + (1-\lambda)x^* \in \text{dom}(f)$ be a neighbor of x^*.
Because f is convex:
$\forall \lambda \in [0,1] : f(z) \leq \lambda f(x) + (1-\lambda)f(x^*)$
$f(x^*) \leq \lambda f(x) + (1-\lambda)f(x^*)$
$f(x^*) \leq f(x)$ Which implies x is a global minimum.

3 Convex or Concave

a: $f(x) = e^x - 1$ is strictly convex.
Proof:

$$f(x) = e^x - 1$$

$$f'(x) = e^x$$
$$f''(x) = e^x$$

By the second order conditions for convexity. this function is strictly convex as it is the case the second derivative is strictly positive.

b: $f(x_1, x_2) = x_1 x_2$ is strictly convex on $\mathbf{R}^2_{++}$
Proof:

$$\nabla f = \begin{vmatrix} x_2 & x_1 \end{vmatrix}$$
$$H_f = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$$

Which is PSD as the $x^T HX$ is strictly greater than 0 for all members in the domain.

c:
$f(w) = \|Xw - y\|_2^2 + \lambda \|w\|_2^2 : \lambda > 0$ is strictly convex.
Proof:

$$\|Xw - y\|_2^2 + \lambda \|w\|_2^2$$
$$= \langle Xw - y, Xw - y\rangle + \lambda \langle w, w\rangle$$
$$= \langle Xw, Xw\rangle - 2\langle Xw, y\rangle + \langle y, y\rangle + \lambda \langle w, w\rangle$$
$$= w^T X^T Xw - 2w^T X^T y + y^T y + \lambda w^T w$$
$$\nabla w = 2X^T Xw - 2X^T y + 2\lambda w$$
$$H_w = 2X^T X + 2\lambda I_n$$

$\forall x \in dom(f) x^T H_w x \geq 0$ therefore the function is strictly convex as it is guaranteed to be greater than 0.
e:

$$f(\mu, \sigma) = nlog\left(\frac{1}{\sqrt{2\pi\sigma}}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

Viewing as a function of just the mean, i.e n log

$$g(\mu) = n\log\left(\frac{1}{\sqrt{2\pi\sigma}}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$g(\mu)$ is strictly concave.
Proof: consider $-g(\mu)$

$$-g'(\mu) = \frac{1}{\sqrt{2\pi\sigma}}\sum_{i=1}^{n} -2x_i + 2\mu$$

Because the derivative of the negative of the function is strictly increasing along the domain of μ, this shows that function is strictly concave
.

Consider

$$-g(z) = -n\log\left(\frac{\sqrt{z}}{\sqrt{2\pi}}\right) + \frac{z}{2}\sum_{i=1}^{n}(x_i-\mu)^2$$

We know that the sums of convex functions are convex.
Let $f_1(z) = \frac{z}{2}\sum_{i=1}^{n}(x_i-\mu)^2$
Let $\lambda \in [0,1]$

$$f_1(\lambda z_1 + (1-\lambda)z_2) = \frac{\lambda z_1 + (1-\lambda)z_2)}{2}\sum_{i=1}^{n}(x_i-\mu)^2$$
$$= \frac{\lambda z_1}{2}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{1-\lambda z_2}{2}\sum_{i=1}^{n}(x_i-\mu)^2$$
$$= f_1(\lambda z_1) + f_1((1-\lambda)z_2)$$
$$\leq f_1(\lambda z_1) + f_1((1-\lambda)z_2)$$

Which shows f_1 is convex.
Let $f_2(z) = -n\log\left(\frac{\sqrt{z}}{\sqrt{2\pi}}\right)$

$$f_2'(z) = -n\frac{\sqrt{2\pi}}{\sqrt{z}}\left(\frac{\sqrt{z}}{\sqrt{2\pi}}\right)' = -\frac{n}{z}$$

This function is strictly increasing along the domain of z, which implies $g(z)$ is strictly concave as the negative of the function is strictly convex.

let $g(\mu,\lambda) = -f(\mu,\lambda)$

$$= -nlog\left(\frac{1}{\sqrt{2\pi}\sigma}\right) + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

If we define $z = \frac{1}{\sigma^2}$

$$\nabla g = \begin{bmatrix} \frac{\partial g}{\partial \mu} & \frac{\partial g}{\partial z} \end{bmatrix}^T$$
$$= \begin{bmatrix} (\frac{1}{\sqrt{2\pi}\sigma}\sum_{i=1}^{n} -2x_i + 2\mu) & -\frac{n}{z} \end{bmatrix}^T$$

By the first order conditions:
A function f is convex iff:

$$\forall x, y \in \mathrm{dom} f, f(y) \geq f(x) + \nabla f(x)^T(y-x)$$

Let $y = (\mu_y, z_y)$, and let $x = (\mu_x, z_x)$

$$f(y) = -n\log\frac{\sqrt{z_y}}{\sqrt{2\pi}} + \frac{z_y}{2}\sum_{i=1}^{n}(x_i-\mu_y)^2$$

$$f(x)+\nabla f(x)^T(y-x) = -n\log\frac{\sqrt{z_x}}{\sqrt{2\pi}} + \frac{z_x}{2}\sum_{i=1}^{n}(x_i-\mu_x)^2 + (\mu_y-\mu_x)(\frac{\sqrt{z_x}}{\sqrt{2\pi}}\sum_{i=1}^{n} -2x_i+2\mu_x) - (z_y-z_x)\frac{n}{z_x}$$

WLOG, let $\mu_y > \mu_y$ and let $z_y = z_x$:

$$f(x) + \nabla f(x)^T(y-x) = -n\log\frac{\sqrt{z_x}}{\sqrt{2\pi}} + \frac{z_x}{2}\sum_{i=1}^{n}(x_i-\mu_x)^2 + (\mu_y-\mu_x)(\frac{\sqrt{z_x}}{\sqrt{2\pi}}\sum_{i=1}^{n}-2x_i+2\mu_x)$$
$$= f(y) + (\mu_y-\mu_x)(\frac{\sqrt{z_x}}{\sqrt{2\pi}}\sum_{i=1}^{n}-2x_i+2\mu_x)$$

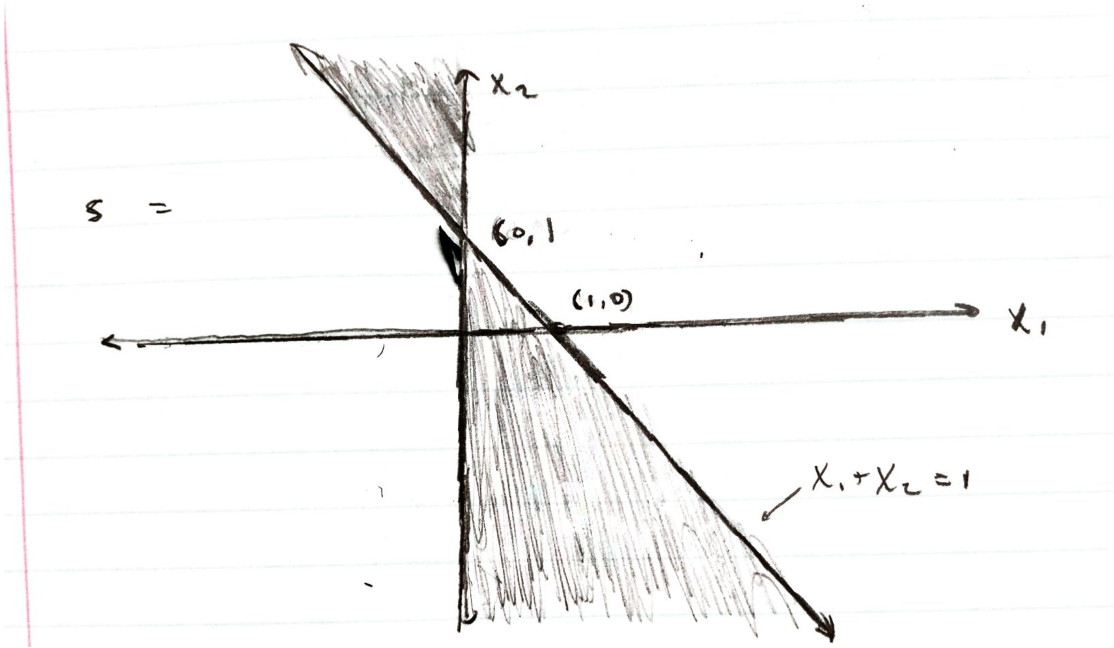
Because we want to show the first order conditions hold, recall we want this to remain true.

$$f(y) \geq f(x) + \nabla f(x)^T(y-x)$$
$$f(y) \geq f(y) + (\mu_y-\mu_x)(\frac{\sqrt{z_x}}{\sqrt{2\pi}}\sum_{i=1}^{n}-2x_i+2\mu_x)$$
$$0 \geq (\mu_y-\mu_x)(\frac{\sqrt{z_x}}{\sqrt{2\pi}}\sum_{i=1}^{n}-2x_i+2\mu_x)$$

Because this inequality is not guaranteed to remain true, ('it depends on the values of x'), this function is not jointly convex.

4 Quadratic Inequalities

$$s_1 = \{x \in \mathbf{R}^2 : x_1 \geq -x_2+1, x_1 \leq 0\}$$
$$s_2 = \{x \in \mathbf{R}^2 : x_1 \leq -x_2+1, x_1 \geq 0\}$$
$$s = s_1 \cup s_2$$



a:

As one can see, this set is not convex as you can draw a line in between two points in the set and the line might leave the set. As an example, consider the points $(1,0)$ and $(0,2)$.

b: Credit to Armin Sukha:
We can describe the solution set S as follows:

$$S_1 \cup S_2 = S = \{x : x_1(-x_2+1) \leq 0\}$$
$$= \{x : -x_1 x_2 + x_1 \leq 0\}$$

Proof:
Case 1: Let $x \in S_1 / S_2$

$$x_1 \leq 0, (-x_2+1) \geq 0$$
$$x_1(-x_2+1) \leq 0$$

Case 2: Let $x \in S_2 / S_1$

$$x_1 \geq 0, (-x_2+1) \leq 0$$
$$x_1(-x_2+1) \leq 0$$
$$\forall x \in S : x_1(-x_2+1) \leq 0$$

given

$$A = -\begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$$
$$x^T A x = -x_1 x_2$$

Similarily, we can define $b = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$, such that $2b^T x = x_1$, we can then just set $c = 0$.
c:
Intuitively, we know that the convex hull of the set is the smallest convex set that contains all members in the set indicated. This is all of $\mathbf{R}^2$. We formally describe the convex set as follows:

$$\text{Cov(S)}\{x \in \mathbf{R}^2 : \frac{1}{2}x_1 + \frac{1}{2}x_2\}$$

5 About General Optimization

Consider the following optimization problem:

$$g^* = \min_{x \in \mathbf{R}^n} f_0(x)$$
$$\text{subject to } : f_i(x) \leq 0, i = 1, ..., m$$
$$Ax = b$$

a: True:
Consider the following optimization problem:

$$t^* = \min_{x\in\mathbf{R}^n, t\in\mathbf{R}} t$$

$$\text{subject to }: f_i(x) \le 0, i = 1, ..., m$$

$$Ax = b$$

$$f_0(x) \le t$$

Proof:
Let t^* be a minimum of the following reformulation of its problem. Because t^* is in the feasible set

$$t^* \in \{x | Ax = b, f_i(x) \le 0, i = 1, ..., m : f_0(x) \le x\}$$

$$t^* \le g^*$$

Because satisfying the set of constraints in the second problem satisfies the set of constraints in the first problem, the feasible set of the first problem is a super-set of the second set.

$$\to g^* \le t^*$$

Therefore $g^* = t^*$
b: Consider the following reformulation:
Define $\mathbf{1}\{0 \text{ if } x \in \mathscr{X} : \infty \text{ otherwise }\}$

$$g^* = \min_x f_0(x) + \mathbf{1}(x)$$

Proof:
Let p^* be a solution of the original problem, because $x \in \mathscr{X}$, this implies that $p^* = g^*$ as the indicator function is equal to zero.

Let g^* be a solution of the reformulation:
if $g^* = \infty \to \mathscr{X} = \emptyset \to p^* = \infty$.
Else $g^* = p^*$ as the indicator will be set to zero.

c: Consider the following reformulation:

$$g^* = \min t$$

$$\exists f(x) \ge t$$

Proof:
let g^* be a solution to the reformulation, and for the sake of contradiction, assume $g^* \ge p^*$.

$$\exists x \in \mathscr{X} : f(x) \le g^*$$

Which implies g^* is not optimal, because we can still satisfy the constraints of the problem by setting $g^* = f(x)$

d:
Let $f(x) = \{x^2 + 2x, x \leq 0 : x^2 - 3x,\ \text{otherwise}\}$
Consider the two following minimization problem:

$$\min_x f(x)$$

$$\text{s.t } x < 0$$

Here, the constraint is slack as the minimum of the function is found at $x = -1$. However, removing the constraint, we achieve a lower minimum value at $x = -1.5$