

Homework 2

Toffalini Leonardo

Exercise 1

Show that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if and only if $g(t) := f(ty + (1-t)x)$ is convex for any $x, y \in \text{dom}(f)$.

Solution

By definition f is convex if and only if $f(ty + (1-t)x) \leq tf(y) + (1-t)f(x)$ for any $t \in [0, 1]$. We need to prove that the previous is equivalent to the following:

$$g(\lambda t_1 + (1-\lambda)t_2) \leq \lambda g(t_1) + (1-\lambda)g(t_2).$$

Expanding the above equation by substituting for the definition of g we get the following:

$$\begin{aligned} g(\lambda t_1 + (1-\lambda)t_2) &= f((\lambda t_1 + (1-\lambda)t_2)y + (1 - (\lambda t_1 + (1-\lambda)t_2))x) \\ &= f(\lambda t_1 y + (1-\lambda t_1)x + (1-\lambda)t_2 y + (1-(1-\lambda)t_2)x) \\ &= f(\lambda(t_1 y + (1-t_1)x) + (1-\lambda)(t_2 y + (1-t_2)x)) \\ &\leq \lambda f(t_1 y + (1-t_1)x) + (1-\lambda)f(t_2 y + (1-t_2)x) \\ &= \lambda g(t_1) + (1-\lambda)g(t_2) \end{aligned}$$

Exercise 2

Prove that

1. e^{ax} is convex on \mathbb{R} for any $a \in \mathbb{R}$,
2. x^a is convex on $\mathbb{R}_{>0}$ when $a \geq 1$ or $a \leq 0$, and is concave when $0 < a < 1$,
3. $\log(x)$ is concave on $\mathbb{R}_{>0}$,
4. $x \log(x)$ is convex on $\mathbb{R}_{>0}$.

Solution

1. e^{ax} is convex for any $a \in \mathbb{R}$

Since $e^{ax} \in C^\infty(\mathbb{R})$ for any $a \in \mathbb{R}$, we can just show that its second derivative is positive. The following is the second derivative:

$$(e^{ax})'' = a^2 e^{ax}.$$

We can clearly see that it is nonnegative, since $a^2 \geq 0, \forall a \in \mathbb{R}$, and $e^c > 0, \forall c \in \mathbb{R}$.

Solution

2. x^a is convex on $\mathbb{R}_{>0}$ when $a \geq 1$ or $a \leq 0$, and is concave when $0 < a < 1$

Since $x^a \in C^2$ for any value $a \in \mathbb{R}$, we just need to verify that in the given intervals it is convex and concave, respectively. The following is the second derivative:

$$(x^a)'' = (ax^{a-1})' = a(a-1)x^{a-2}.$$

Since $x^c \geq 0$ holds for all $x > 0$, we just need to check for which values of a is $a(a-1)$ negative or positive. With some elementary arithmetic we can verify that $a(a-1)$ is negative if and only if $a \in (0, 1)$. Thus, the function is concave on said interval, and is convex elsewhere.

To tidy up some negligence from the previous argument, we still need to take care of the cases where $a = 1$ and $a = 0$, since in these cases the power rule does not hold for $(x^0)'$. Luckily it is trivial to check that the linear function x and the constant function $x^0 = 1$ are both convex.

Solution

3. $\log(x)$ is concave on $\mathbb{R}_{>0}$

Once again, because $\log x \in C^2\mathbb{R}$ we just need to check the sign of $(\log x)''$.

$$(\log x)'' = \left(\frac{1}{x}\right)' = -\frac{1}{x^2} < 0 \quad \forall x \in \mathbb{R}_{>0}$$

Solution

4. $x \log(x)$ is convex on $\mathbb{R}_{>0}$

$$(x \log x)'' = \left(\log x + x \frac{1}{x}\right)' = \frac{1}{x} > 0 \quad \forall x \in \mathbb{R}_{>0}$$

Exercise 3

Show the following.

1. (Nonnegative weighted sum) If $w_i \geq 0$ and f_i is convex for every i , then $f = \sum_i w_i f_i$ is convex.
2. (Affine mapping) If f is convex, then $g(x) = f(Ax + b)$ is convex.
3. (Pointwise maximum) If f_i is convex for every i , then $f(x) = \max_i \{f_i(x)\}$ is convex.
4. (Composition) If g is convex and h is convex and non-decreasing, then $f(x) = h(g(x))$ is convex.

Solution

Nonnegative weighted sum

$$\begin{aligned}
f(tx + (1-t)y) &= \sum w_i f_i(tx + (1-t)y) \leq \sum w_i (tf_i(x) + (1-t)f_i(y)) \\
&= t \sum w_i f_i(x) + (1-t) \sum w_i f_i(y) = tf(x) + (1-t)f(y).
\end{aligned}$$

Solution

Affine mapping

$$\begin{aligned} g(tx + (1-t)y) &= f(A(tx + (1-t)y) + b) = f(tAx + (1-t)Ay + b) \leq \\ &\leq tf(Ax + b) + (1-t)f(Ay + b) = tg(x) + (1-t)g(y). \end{aligned}$$

Solution

Pointwise maximum

$$f(tx + (1-t)y) = \left(\max_i f_i \right)(tx + (1-t)y)$$

Let f_k be the convex function at which the maximum is obtained.

$$\begin{aligned} \left(\max_i f_i \right)(tx + (1-t)y) &= f_k(tx + (1-t)y) \leq tf_k(x) + (1-t)f_k(y) \leq \\ &\leq t \left(\max_i f_i \right)(x) + (1-t) \left(\max_i f_i \right)(y) = tf(x) + (1-t)f(y). \end{aligned}$$

Solution

Composition

$$f(tx + (1-t)y) = h(g(tx + (1-t)y))$$

Since g is convex the inner sum can be broken into something bigger, and since h is non-decreasing, if we put something larger inside the body the value will not get smaller.

$$\leq h(tg(x) + (1-t)g(y)) \leq th(g(y)) + (1-t)h(g(y)) = tf(x) + (1-t)f(y).$$

Exercise 4

Prove that for an arbitrary set $S \subseteq \mathbb{R}^n$, the polar set

$$S^* = \{y \in \mathbb{R}^n \mid y^T x \leq 1 \quad \forall x \in S\}$$

is convex.

Solution

By the definition of a convex set, we have to prove the following:

$$\forall x, y \in S^*, t \in [0, 1] : \quad tx + (1-t)y \in S^*.$$

Since $x, y \in S^*$ we have $x \cdot z \leq 1$ and $y \cdot z \leq 1$ for all $z \in S$.

$$(tx + (1-t)y) \cdot z = tx \cdot z + (1-t)y \cdot z \leq t + (1-t) = 1.$$

Exercise 5

Let us consider the following functions:

$$f_1(x, y) = \frac{1}{2}x^2 + \frac{7}{2}y^2$$

$$f_2(x, y) = 100(y - x^2)^2 + (1 - x)^2 \quad (\text{Rosenbrock's function})$$

$$f_3(x, y) = \frac{1}{2}x^2 + x \cos y.$$

1. Calculate the gradient of the functions.
2. Determine the global minimum of the functions.
3. Are these function convex?

Solution

1.

1.1

$$\nabla f_1(x, y) = \begin{bmatrix} \partial_x f_1(x, y) \\ \partial_y f_1(x, y) \end{bmatrix} = \begin{bmatrix} x \\ 7y \end{bmatrix}$$

2.1 The global minimum value of f_1 is 0, which is attained at the point $(0, 0)$.

3.1 The function is in fact convex, because it's Hessian is $\begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$, which is positive definite.

Solution

2.

2.1

$$\begin{aligned} \nabla f_2(x, y) &= \begin{bmatrix} 200(y - x^2) \cdot (-2x) + 2(1 - x^2) \cdot (-1) \\ 200(y - x^2) \end{bmatrix} = \\ &= \begin{bmatrix} -400x(y - x^2) - 2(1 - x^2) \\ 200(y - x^2) \end{bmatrix} = \\ &= \begin{bmatrix} 400x^3 - 400xy + 2x - 2 \\ 200(y - x^2) \end{bmatrix} \end{aligned}$$

2.2 The global minimum value is 0, which is attained at the point $(1, 1)$.

2.3 The hessian of the Rosenbrock function is the following:

$$H = \begin{bmatrix} 1200x^2 - 400y + 2 & -400x \\ -400x & 200 \end{bmatrix}.$$

The hessian at point $(x, y) = (1, 1.5)$ is the following:

$$H = \begin{bmatrix} 602 & -400 \\ -400 & 200 \end{bmatrix}.$$

At this point the full determinant is $602 \cdot 200 - 400 \cdot 400 = -39600 < 0$. Thus, the hessian is not PSD making it non convex.

Solution

3.

3.1

$$\nabla f_3(x, y) = \begin{bmatrix} x + \cos y \\ -x \sin y \end{bmatrix}$$

3.2 The minimum value of -1 is attained at the point $(x, y) = (-1, 0)$.

3.3 The hessian of the function is the following:

$$H = \begin{bmatrix} 1 & -\sin y \\ -\sin y & -x \cos y \end{bmatrix}$$

The full determinant of the matrix is $-x \cos y - \sin^2 y$, which at the point $(1, 0)$ is -1 , making the hessian not PSD, thus the function is not convex.

Exercise 6

Prove that for an arbitrary function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the conjugate function

$$f^*(y) = \sup\{y^T x - f(x) \mid x \in \text{dom}(f)\}$$

is convex.

Solution

$$\begin{aligned} f^*(ty + (1-t)z) &= \sup_{x \in \text{dom}(f)} \{(ty + (1-t)z) \cdot x - f(x)\} \\ &= \sup_{x \in \text{dom}(f)} \{t(y \cdot x) + (1-t)(z \cdot x) - f(x)\} \\ &= \sup_{x \in \text{dom}(f)} \{t(y \cdot x) + (1-t)(z \cdot x) - (tf(x) + (1-t)f(x))\} \\ &= \sup_{x \in \text{dom}(f)} \{t(y \cdot x - f(x)) + (1-t)(z \cdot x - f(x))\} \\ &\leq \sup_{x \in \text{dom}(f)} \{t(y \cdot x - f(x))\} + \sup_{x \in \text{dom}(f)} \{(1-t)(z \cdot x - f(x))\} \\ &= t \sup_{x \in \text{dom}(f)} \{y \cdot x - f(x)\} + (1-t) \sup_{x \in \text{dom}(f)} \{z \cdot x - f(x)\} \\ &= tf^*(y) + (1-t)f^*(z) \end{aligned}$$

Exercise 7

Consider the optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & x^2 + 2x + 4 \\ \text{s.t.} \quad & x^2 - 4x \leq -3 \end{aligned}$$

1. Solve this problem, i.e., find the optimal solution.
2. Derive the Lagrangian dual problem.
3. Prove, that weak duality holds by computing the maximum of the dual problem. Is Slater's condition satisfied? Does strong duality hold?

Solution

1. Since the objective function is a convex quadratic function, the optimum will be attained at either the point where the first derivative vanishes, or on the boundary.

The objective function's derivative is $2x + 2$, which has a root at -1 . However, we still need to check if the defining condition is met for $x = -1$.

$$(-1)^2 - 4(-1) = 1 + 4 = 5 \not\leq -3.$$

So $x = -1$ is not the solution. Thus, we need to investigate the boundary of the condition.

We need to check for what x does the following hold:

$$x^2 - 4x + 3 \leq 0.$$

This inequality holds for $x \in [1, 3]$. From this it is easily verifiable, that the solution is found at $x = 1$ with the value of 7.

Solution

2. The Lagrangian dual function is the following:

$$\begin{aligned} g(\lambda, \mu) &= \inf_x L(x, \lambda, \mu) \\ &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \mu_i h_i(x) \right) \end{aligned}$$

In our case $f_0 = x^2 + 2x + 4$, and $f_1 = x^2 - 4x + 3$, and there are not other f_i or h_i .

$$\begin{aligned} g(\lambda, \mu) &= \inf_x (x^2 + 2x + 4 + \lambda(x^2 - 4x + 3)) \\ &= \inf_x ((1 + \lambda)x^2 + (2 - 4\lambda)x + 4 + 3\lambda) \end{aligned}$$

Thus the Lagrangian dual problem is the following:

$$\begin{aligned} &\text{maximize} \quad \inf_x ((1 + \lambda)x^2 + (2 - 4\lambda)x + 4 + 3\lambda) \\ &\text{subject to} \quad \lambda \geq 0 \end{aligned}$$

Solution

3. Let's solve the dual problem. Since the dual function is always concave, independent of the convexity of the primal function, we just need to check where the first derivative vanishes.

$$\begin{aligned} 0 &= ((1 + \lambda)x^2 + (2 - 4\lambda)x + 4 + 3\lambda)' = 2(1 + \lambda)x + 2 - 4\lambda \\ 0 &= 4 - 2\lambda \implies \lambda = 2. \end{aligned}$$

At $\lambda = 2$ the maximal value of the dual function is

$$g(2) = \inf_x (3x^2 - 6x + 10) = 7.$$

From this we see that not only weak duality, but strong duality holds, as $d^* = p^* = 7$.

We can also see that Slater's conditions are met, since f_1 is convex and there exists a point $x \in (1, 3)$, which strictly satisfies the condition $f_1(x) < 0$.

Exercise 8

Given a convex, differentiable function $F : K \rightarrow \mathbb{R}$ over a convex subset K of \mathbb{R}^n , the *Bregman divergence* of $x, y \in K$ is defined as $D_F(x, y) = F(y) - F(x) - \langle \nabla F(x), y - x \rangle$. Prove that $D_F(x, y) \geq 0$. Define a function F for which $D_F(x, y) = \|x - y\|_2^2$.

Solution

Since F is convex and differentiable, the first order characterization holds:

$$F(y) \geq F(x) + \langle \nabla F(x), y - x \rangle.$$

Rearranging the inequality we get the following:

$$F(y) - F(x) - \langle \nabla F(x), y - x \rangle \geq 0.$$

The left hand side of the inequality is precisely the definition of $D_F(x, y)$, thus it is nonnegative.

In the following we show that for $F(x) = \|x\|^2$, we get $D_F(x, y) = \|x - y\|^2$:

$$\begin{aligned} F(x) = \|x\|^2 &\implies \nabla F(x) = \nabla(x^T I x) = 2x \\ D_F(x, y) &= \|y\|^2 - \|x\|^2 - \langle 2x, y - x \rangle \\ &= \|y\|^2 - \|x\|^2 - \langle 2x, y \rangle + \langle 2x, x \rangle \\ &= \|y\|^2 - \|x\|^2 - 2\langle x, y \rangle + 2\|x\|^2 \\ &= \|x\|^2 - 2\langle x, y \rangle + \|y\|^2 \\ &= \langle x - y, x - y \rangle = \|x - y\|^2 \end{aligned}$$