Name: John Min, UNI: jcm2199

COMS 4772 Homework Set 1

Try to do all problems efficiently, but show key steps in your work.

Multi-variable Calculus Review Problems

(1) Find the global minimizers, if they exist, for the following functions.

(a) $f(x) = x_1^2 - 4x_1 + 2x_2^2 + 7$

Solution:
$\frac{\partial f}{\partial x_1} = 2x_1 - 4, \frac{\partial f}{\partial x_2} = 4x_2$
$\frac{\partial^2 f}{\partial x_1^2} = 2, \frac{\partial^2 f}{\partial x_1^2}$
$2x_1^* - 4 = 0, 4x_2^* = 0 \Rightarrow x^* = (2, 0)$.

(b) $f(x) = e^{-\|x\|^2}$

Solution:
A global minimizer does not exist as this function is concave and for x^* : $\nabla f(x^*) = 0$, we are at a global maximum ($f''(0) < 0$).

(c) $f(x) = x_1^2 - 2x_1x_2 + \frac{1}{3}x_2^3 - 4x_2$

Solution:
$\frac{\partial f}{\partial x_1} = 2x_1 - x_2 := 0 \Rightarrow x_2 = 2x_1$
$\frac{\partial f}{\partial x_2} = -x_1 + x_2^2 - 4 := 0 \Rightarrow 4x_1^2 - x_1 - 4 = 0 \Rightarrow x_1^* = \pm\frac{\sqrt{17}}{8}$
$\frac{\partial^2 f}{\partial x_1^2} = 2, \frac{\partial^2 f}{\partial x_2^2} = 2x_2$
For our global minimum, we need $x_2 > 0$; hence, we choose $x^* = (\frac{\sqrt{17}}{8}, \frac{\sqrt{17}}{4})$.

(d) $f(x) = (2x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - 1)^2$

Solution:
$f(x) = 4x_1^2 - 4x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2 - 2x_3 + 1$
$\frac{\partial f}{\partial x_1} = 8x_1 - 4x_2 := 0 \Rightarrow 2x_1 = x_2$
$\frac{\partial f}{\partial x_2} = -4x_1 + 4x_2 - 2x_3 := 0 \Rightarrow x_2 - x_3 = 0 \Rightarrow x_2 = x_3$
$\frac{\partial f}{\partial x_3} = -2x_2 + 4x_3 - 2 := 0 \Rightarrow 2x_3 - 2 = 0 \Rightarrow x_3 = 1$
$\frac{\partial^2 f}{\partial x_1^2} = 8, \frac{\partial^2 f}{\partial x_2^2} = 4, \frac{\partial^2 f}{\partial x_3^2} = 4$
$x^* = (\frac{1}{2}, 1, 1)$

(e) $f(x) = x_1^4 + 16x_1x_2 + x_2^8$

Solution:
$\frac{\partial f}{\partial x_1} = 4x_1^3 + 16x_2 := 0 \Rightarrow x_1^3 + 4x_2 = 0 \Rightarrow x_1^3 = -4x_2$

$\frac{\partial f}{\partial x_2} = 16x_1 + 8x_2^7 = 0 \Rightarrow 2x_1 + x_2^7 = 0 \Rightarrow 2x_1 = -x_2^7$
$\frac{\partial^2 f}{\partial x_1^2} = 12x_1^2, \frac{\partial^2 f}{\partial x_2^2} = 56x_2^6$. Hence, the Hessian $\nabla^2 f$ is semi-positive definite.
Solve the partial equations and $x^* = (-2^{\frac{3}{4}}, 2^{\frac{1}{4}})$.

(f) $f(x) = \sum_{j=1}^{n-1} 10^j (x_j - x_{j+1}^2)^2$ (The Rosenbrock function)

Solution:
Since this function is a sum of squares, the global minimum is 0. In this case, we see that if $x^* = (c, c^{1/2}, \ldots, c^{1/2^{n-1}})$, for $c \geq 0$ such that $x_i = c^{1/2^{i-1}}$ for $i = 1, \ldots, n$, then, we are at a global min.

(2) When f is continuously differentiable at $x \in \mathbb{R}^n$, then $\nabla f(x)$ is easily computed as the vector of partial derivatives of f at x. Compute the gradient of the following functions.

(a) $f(x) = x_1^3 + x_2^3 - 3x_1 - 15x_2 + 25$

Solution:
$\nabla f(x) = [3x_1^2 - 3, 3x_2^2 - 15]$

(b) $f(x) = x_1^2 + x_2^2 - \sin(x_1 x_2)$

Solution:
$\nabla f(x) = [2x_1 - cos(x_1 \cdot x_2) \cdot x_2, 2x_2 - cos(x_1 \cdot x_2) \cdot x_1]$

(c) $f(x) = \|x\|^2 = \sum_{j=1}^{n} x_j^2$

Solution:
$\nabla f(x) = 2 \cdot [x_1, x_2, \ldots, x_n]$

(d) $f(x) = e^{\|x\|^2}$

Solution:
$$\nabla f(x) = 2f(x) \cdot [x_1, x_2, \ldots, x_n] = 2 \cdot \exp\{\sum_{j=1}^{n} x_j^2\} \cdot [x_1, x_2, \ldots, x_n]$$

(e) $f(x) = x_1 x_2 x_3 \cdots x_n$

Solution:
$\nabla f(x) = f(x) \cdot [x_1, x_2, \ldots, x_n]^{-1} = x_1 x_2 x_3 \cdots x_n \cdot [\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}]$

(f) $f(x) = -\log(x_1 x_2 x_3 \cdots x_n)$ for $x_j > 0,\ j = 1, \ldots n$.

Solution:
$\nabla f(x)$ is a vector of the following partial:

$$\frac{\partial f}{\partial x_i} = \frac{-1}{log(\prod_i x_i)} \cdot \prod_{i \neq j}(x_i)$$
$$\nabla f(x) = [\frac{-1}{log(\prod_i x_i)} \cdot \prod_{i \neq 1}(x_i), \ldots, \frac{-1}{log(\prod_i x_i)} \cdot \prod_{i \neq n}(x_i)]$$

(3) Let $\mathbb{R}^{n \times n}$ denote the set of real $n \times n$ square matrices A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be twice differentiable at a point $x \in \mathbb{R}^n$ if is differentiable at x and there is a matrix $H \in \mathbb{R}^{n \times n}$ such that

$$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2}(y - x)^T H(y - x) + o(\|y - x\|^2).$$

The matrix H is called the Hessian of f at x and is denoted $\nabla^2 f(x)$. Note that, when defined, the relation $x \mapsto \nabla^2 f(x)$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^{n \times n}$, i.e. $\nabla^2 f : \mathbb{R}^n \to \mathbb{R}^n$. We say that f is twice continuously differentiable at $x \in \mathbb{R}^n$ if the mapping $\nabla^2 f$ is continuous at x. It can be shown that if f is twice continuously differentiable at a point $x \in \mathbb{R}^n$, then the matrix $\nabla^2 f(x)$ is symmetric, i.e. $\nabla^2 f(x) = \nabla^2 f(x)^T$, in which case $\nabla^2 f(x)$ is the matrix of second partial derivatives of f at x:

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

Compute the Hessian of the functions given in problem (2) above.

(a) $f(x) = x_1^3 + x_2^3 - 3x_1 - 15x_2 + 25$

Solution:
$$\nabla^2 f(x) = \begin{bmatrix} 6x_1 & 0 \\ 0 & 6x_2 \end{bmatrix}$$

(b) $f(x) = x_1^2 + x_2^2 - \sin(x_1 x_2)$

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

(c) $f(x) = \|x\|^2 = \sum_{j=1}^n x_j^2$

Solution:
$$\nabla^2 f(x) = 2 \cdot \text{diag}(e) = \begin{bmatrix} 2 & & & \\ & 2 & & \\ & & \ddots & \\ & & & 2 \end{bmatrix}$$

(d) $f(x) = e^{\|x\|^2}$

Solution:
$$\nabla^2 f(x) = 4 \cdot \exp\{\sum_{j=1}^{n} x_j^2\} \cdot x^\top x$$

(e) $f(x) = x_1 x_2 x_3 \cdots x_n$

Solution:
$$\nabla^2 f(x) = \frac{f(x)}{x^\top x}$$

(f) $f(x) = -\log(x_1 x_2 x_3 \cdots x_n)$ for $x_j > 0,\ j = 1, \ldots n$.

Solution:
$$\frac{\partial f}{\partial x_i} = \frac{-1}{\prod_i x_i} \prod_{i \neq j} x_i = -\frac{1}{x_i}$$
$$\nabla^2 f(x) = \mathrm{diag}(x_i^{-2})$$

Working with Convex Functions

(4) Show that each of the following functions is convex.

(a) $f(x) = \begin{cases} 12 & \text{if} \quad \|Ax - b\|_1 \le \pi \\ \infty & \text{otherwise} \end{cases}$

Solution:
$f(x) = 12 + \delta(x|C)$ where $C = \{x : \|Ax - b\|_1 \le \pi\}$.
As stated in the notes, all norms can be written as support functions and thus, all norms are convex. Thus, C is convex and the convex indicator function is convex on a convex domain.

(b) $f(x) = e^{-x}$

Solution:
f is a twice differentiable continuous function
$f'(x) = -e^{-x}$
$f''(x) = e^{-x} > 0\ \forall x$.

(c) $f(x_1, x_2, \ldots, x_n) = e^{-(x_1 + x_2 + \cdots + x_n)}$

Solution:
f is again a twice differentiable continuous function.
$\nabla f(x) = -f(x), \nabla^2 f(x) = f(x)$.
Since $f(x) > 0$, f is clearly convex.

(d) $f(x) = \|x\|$ (Show for any norm, don't assume this is a particular norm).

Solution:

$f(\lambda x + (1-\lambda)y) \leq f(\lambda x) + f((1-\lambda)y)$ (triangle inequality)

$= \lambda \cdot f(x) + (1-\lambda) \cdot f(y)$ (homogeneity of norms)

(e) $f(x) = \log\left(\sum_i e^{x_i}\right)$

Solution:
$\dfrac{\partial f}{\partial x_i} = \dfrac{e^{x_i}}{\sum_i e^{x_i}}$
$\dfrac{\partial^2 f}{\partial x_i^2} = e^{x_i} e^{x_j} (\sum_k e^{x_k})^{-2}$
$\nabla f(x) = \dfrac{1}{e^\top z}\mathrm{diag}(z)$ where $z = e^X (z_i = e^{x_i})$.

$\nabla^2 f(x) = \frac{1}{e^\top z}\mathrm{diag}(z) - \frac{1}{(e^\top z)^2} zz^\top$.

We need to show that $\nabla^2 f(x)$ is semi-positive definite:

$$\begin{aligned} x^\top \nabla_f^2 x &= \frac{1}{e^\top z}[x^\top \mathrm{diag}(z) x \cdot e^\top z - (z^\top x)^\top)(z^\top x)] \\ &= \frac{1}{e^\top z}[(\sum_i x_i^2 z_i) \cdot (\sum_i z_i) - (\sum_i x_i z_i)^2] \end{aligned}$$

The denominator is clearly positive and by Cauchy-Schwarz, it is clear that the numerator is non-negative. Hence, the Hessian is S.P.D.

(5) If $f_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, 2$ are convex, show that $f(x) = \max\{f_1(x),\ f_2(x)\}$ is a convex function.

Solution:

Let $z = \lambda x + (1-\lambda)y$.

$$\begin{aligned} f(z) &= \max\{f_1(z), f_2(z)\} \\ &= \max\{f_1(\lambda x + (1-\lambda)y), f_2(\lambda x + (1-\lambda)y)\} \\ &\leq \max\{\lambda f_1(x) + (1-\lambda) f_1(y), \lambda f_2(x) + (1-\lambda) f_2(y)\} \text{ by the convexity of } f_i \\ &\leq \max\{\lambda \cdot \max\{f_1(x),\ f_2(x)\} + (1-\lambda) \cdot \max\{f_1(x),\ f_2(x)\}\} \\ &= \lambda \cdot \max\{f_1(x),\ f_2(x)\} + (1-\lambda) \cdot \max\{f_1(x),\ f_2(x)\} \\ &= \lambda \cdot f(x) + (1-\lambda) \cdot f(y). \quad \square \end{aligned}$$

(6) Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, and suppose that $f : \mathbb{R}^m \to \mathbb{R}$ is convex. Show that $h(x) = f(Ax + b)$ is convex.

Solution:

Let $z = \lambda x + (1-\lambda)y$.

$$\begin{aligned} h(z) &= f(Az+b) \\ &= f(A[\lambda x + (1-\lambda)y] + b)] \\ &= f(\lambda[Ax+b] + (1-\lambda)[Ay+b]) \\ &\le \lambda \cdot f(Ax+b) + (1-\lambda) \cdot f(Ay+b) \\ &= \lambda \cdot h(x) + (1-\lambda) \cdot h(y). \quad \square \end{aligned}$$

(7) Consider the linear equation

$$Ax = b,$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. When $n < m$ it is often the case that this equation is over-determined in the sense that no solution x exists. In such cases one often attempts to locate a 'best' solution in a least squares sense. That is one solves the *linear least squares problem*

$$\text{(lls)} : \text{minimize } \frac{1}{2}\|Ax - b\|_2^2$$

for x. Define $f : \mathbb{R}^n \to \mathbb{R}$ by

$$f(x) := \frac{1}{2}\|Ax - b\|_2^2.$$

(a) Show that f can be written as a quadratic function, i.e. a function of the form

$$f(x) := \frac{1}{2}x^T Q x - a^T x + \alpha \ .$$

Solution:

$$\begin{aligned} f(x) = \frac{1}{2}\|Ax-b\|_2^2 &= \frac{1}{2}\|Ax-b\|_2\|Ax-b\|_2 \\ &= \frac{1}{2}(Ax-b)^\top(Ax-b) \\ &= \frac{1}{2}\{(Ax)^\top(Ax) - b^\top(Ax) - (Ax)^\top b + b^\top b\} \\ &= \frac{1}{2}\{x^\top A^\top Ax - 2b^\top Ax + b^\top b\} \end{aligned}$$

Set $Q = A^\top A, a = A^\top b, \alpha = \frac{1}{2}b^\top b$.

(b) What are $\nabla f(x)$ and $\nabla^2 f(x)$?

Solution:
$\nabla f(x) = Qx - a = A^\top Ax - A^\top b = A^\top(Ax - b)$
$\nabla^2 f(x) = Q = A^\top A$

(c) Show that $\nabla^2 f(x)$ is positive semi-definite.

Solution:
$\forall z \in \mathbb{R}^n, z^\top \nabla^2 f z = z^\top A^\top A z = (Az)^\top (Az) = \|Az\|_2^2 \geq 0.$
Hence, Q is P.S.D.

(d) * Show that a solution to (lls) must always exist.

Solution:
$\nabla f(x) = 0 : A^\top Ax = A^\top b.$
Since A is a square, symmetric matrix, $\text{rank}(A^\top A) = \text{rank}(A) = \text{rank}(A^\top)$. Therefore, we can find a solution to the system (lls).

(e) * Provide a necessary and sufficient condition on the matrix A (**not on the matrix** $A^T A$) under which (lls) has a unique solution and then display this solution in terms of the data A and b.

Solution:
A unique solution exists if and only if $A^\top Ax = A^\top b$ is unique. We need $x^* = (A^\top A)^{-1} A^\top b$,with $A^\top A$ nonsingular. When $Q = A^\top A$ is nonsingular, hence, invertible, $\|Q^{-1}\| = \{\min_{\|x\|_2=1} \|Ax\|_2\}^{-1} = \dfrac{1}{\sqrt{\lambda_{\min}}}$ where $\lambda_{\min}$ is the smallest eigenvalue of $A^\top A$. Clearly, $\lambda_{\min} > 0$ for nonsingularity. Hence, for A, where $v_{\min}$ is the smallest eigenvalue of A, we need $v_{\min} = (\lambda_{\min})^{1/2}) > 0$. More simply, A should be a full rank matrix.

(8) Consider the functions

$$f(x) = \frac{1}{2} x^T Q x - c^T x$$

and

$$f_t(x) = \frac{1}{2} x^T Q x - c^T x + t\phi(x),$$

where $t > 0$, $Q \in \mathbb{R}^{n \times n}$ is positive semi-definite, $c \in \mathbb{R}^n$, and $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is given by

$$\phi(x) = \begin{cases} -\sum_{i=1}^n \ln x_i & \text{, if } x_i > 0,\ i = 1, 2, \ldots, n, \\ +\infty & \text{, otherwise.} \end{cases}$$

(a) Show that ϕ is a convex function.

Solution:
The log function is concave, hence, the - log function is convex. The sum of convex functions are convex.
$\nabla_x \phi = -\frac{1}{x_i} = -x_i^{-1}$
$\nabla_x^2 \phi = \text{ diag}(x_i^2)^{-1}.$
Clearly, positive definite as $x_i > 0$.

(b) Show that both f and f_t are convex functions.

Solution:
$\nabla f(x) = Qx$
$\nabla^2 f(x) = Q$, which is P.S.D. $\Rightarrow x^\top Qx$ is convex, $-c^\top x$ is linear, hence, convex.
f is convex, ϕ is convex $\Rightarrow f_t$ convex.

(c) Show that the solution to the problem $\min f_t(x)$ always exists and is unique.

$\nabla^2 f_t = \nabla^2 f + \nabla^2 \phi$ is symmetric and positive definite since $\nabla_x^2 \phi$ is symmetric and positive definite and $\nabla^2 f = Q$ is symmetric semi-positive definite.
$\Rightarrow f_t$ is strictly convex.
Therefore, a solution exists and is unique.

Working with Convex Functions, part II.

(9) The conjugate of a convex function is defined by

$$f^*(y) = \sup_x x^T y - f(x).$$

Compute the conjugates of the following functions.

(a) $f(x) = \delta(x|C)$, where C is a convex set.

Solution:
$f^*(x) = \sup_x x^\top y - \delta(x|C) = \sup_{x \in C} x^\top y$ the support function of C.

(b) $f(x) = 3x^2 - 2x$

Solution:
$\frac{\partial f}{\partial x} = 0 : y - 6\bar{x} + 2 = 0 \Rightarrow \bar{x} = \frac{1}{6}(y+2)$

$$\begin{aligned} f^*(y) &= \frac{1}{6}(y+2)y - \frac{1}{12}(y^2 + 4y + 4) + \frac{1}{3}y + \frac{2}{3} \\ &= \frac{1}{12}(y^2 + 4y + 4) \\ &= \frac{1}{12}(y+2)^2 \end{aligned}$$

(c) $f(x) = x \log x$

Solution:
$\frac{\partial f}{\partial x} = 0 : y - \log \bar{x} - 1 = 0 \Leftrightarrow \log \bar{x} = y - 1 \Leftrightarrow \bar{x} = \exp\{y-1\}$

$$\begin{aligned} f^*(y) &= \sup_x < x, y > -x \log x \\ &= y \cdot \exp\{y-1\} - (y-1)\exp\{y-1\} \\ &= \exp\{y-1\} \end{aligned}$$

(d) $f(x) = \log(1 + e^x)$

Solution:
$f^*(y) = \sup_x x^\top y - \log(1 + e^x)$

Differentiate: $(f^*)' = 0 : y - \dfrac{e^x}{1+e^x} = 0$

$x^* = -\log(\dfrac{1}{y} - 1)$

$\Rightarrow f^*(y) = -y \log(\dfrac{1}{y} - 1) - \log(\dfrac{2y-1}{1-y})$

(e) $f(x) = e^{-2x}$.

Solution:
$f^*(y) = \sup_x x^\top y - e^{-x}$

Differentiate: $(f^*)' = 0 : y + e^{-x} = 0 \Rightarrow e^{-x} = -y \Rightarrow x^* = \log(y)$ for $y > 0$.

$$f^*(y) = \begin{cases} y\log y - e^{-\log(y)} = y \log y - \frac{1}{y}, & \text{for } y > 0 \\ 0, & \text{otherwise} \end{cases}$$

(f) $f(x) = \max_{1,\ldots,n} x_i$.

Solution:

$$f^*(x) = \begin{cases} 0, & \text{if } y > 0 \text{ and } \|y\| \leq 1 \\ \infty, & \text{otherwise} \end{cases}$$

If $y \leq 0$, we can choose an arbitrarily large x_i. The only way this conjugate function is bounded is if the norm of y is bounded to be at most that of a unit vector. When $\|y\| = 1$, $\sum_i x_i \cdot y_i \leq \max_i x_i$ holds in equality.

(g) $f(x) = x^p$ for $x > 0$ and $p > 1$.

Solution:
$f^*(y) = \sup_x x^\top y - x^p$
Differentiate $(f^*)' = 0 : y - p\bar{x}^{p-1} = 0 \Rightarrow \bar{x} = [\frac{y}{p}]^{1/(p-1)}$
Substitute $\bar{x}$ in for y:

$$y \cdot (\frac{y}{p})^{1/(p-1)} - (\frac{y}{p})^{p/(p-1)}$$
$$[p-1](\frac{y}{p})^{p/(p-1)} \text{for } y > 0$$

$$f^*(y) = \begin{cases} [p-1](\frac{y}{p})^{p/(p-1)}, & \text{for } y > 0 \\ 0, & \text{otherwise} \end{cases}$$

(10) Conjugates in terms of other conjugates

(a) Define $g(x) = f(x) + c^T x + d$, where f is convex. Express g^* in terms of f^*, c, and d.

Solution:

$$\begin{aligned} g^*(y) &= \sup_x x^\top y - f(x) - c^\top x - d \\ &= \sup_x \{x^\top y - f(x) - c^\top x\} - d \\ &= \sup_x \{< x, y - c > - f(x)\} - d \\ &= f^*(y - c) - d \end{aligned}$$

(b) Define $g(x) = f(Qx + b)$, Q an invertible matrix, and b a vector. Compute f^* in terms of g^*, Q, and b.

Solution:
Let $u = Qx + b$. $x = Q^{-1}(u - b)$.

$$\begin{aligned} g^*(y) &= \sup_u < u, Q^{-\top}y > -f(u) - b^\top Q^{-\top}y \\ &= f^*(Q^{-\top}y) - (Q^{-1}b)^\top y \end{aligned}$$

(c) Let $f(x, z)$ be convex in (x, z) and define $g(z) = \inf_z f(x, z)$. Express g^* in terms of f^*.

Solution:

$$\begin{aligned} g^*(y) &= \sup_x < x, y > - \inf_z f(x, z) \\ &= \sup_{x,z}\{< x, z > -f(x, z)\} \\ &= \sup_z\{\sup_x < x, y > -f(x, z)\} \\ &= \sup_z f^*(y) \end{aligned}$$

(11) Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex, and bounded above on $\mathbb{R}^n$. Show that f is constant.

Solution:
Suppose f is not constant. Then, $\exists x, y$, s.t. $f(x) > f(y)$.
$f(\lambda(x) + (1 - \lambda)y) \leq \lambda \cdot f(x) + (1 - \lambda) \cdot f(y)$.
Let $z = \lambda \cdot x + (1 - \lambda) \cdot y$
By convexity, $\dfrac{f(z) - (1 - \lambda) \cdot f(y)}{\lambda} \leq f(x)$ where $\lambda \in [0, 1]$.
$\dfrac{f(z) - (1 - \lambda) \cdot f(y)}{\lambda} = \dfrac{f(z) - f(y)}{\lambda} + f(y) \leq f(x)$.
As $\lambda \to 0^+$, the first term blows up and it is clear that $f(x)$ is unbounded.

Bonus: Don't work on this problem until you are done with the others.

(12) Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is convex and twice continuously differentiable. Suppose that $\bar{y}$ and $\bar{x}$ are related by $\bar{y} = \nabla f(\bar{x})$, and that the hessian $\nabla^2 f(\bar{x})$ is positive definite.

(a) Show that $\nabla f^*(\bar{y}) = \bar{x}$.

Solution:
$f^*(\bar{y}) = \sup_x x^\top \bar{y} - f(x)$
$\dfrac{\partial f^*}{\partial x} = 0 : \bar{y} - f'(x) = 0 \Rightarrow \bar{y} = \nabla f(x^*) = \nabla f(\bar{x}) \Rightarrow x^* = \bar{x}$

Therefore, $f^*(\bar{y}) = \sup_x \bar{x}^\top \bar{y} - f(\bar{x})$.
Differentiate w.r.t. $\bar{y}$ and we get $\nabla f^*(\bar{y}) = \bar{x}$.

(b) Show that $\nabla^2 f^*(\bar{y}) = \nabla^2 f(\bar{x})^{-1}$.

Solution:
$\partial \bar{y} = \nabla^2 f(\bar{x}) \partial \bar{x} \Rightarrow \frac{\partial \bar{x}}{\partial \bar{y}} = [\nabla^2 f(\bar{x})]^{-1} = \nabla^2 f^*(\bar{y})$