HW1

Tuesday, January 28, 2025 12:52 PM

Exercise 1.15 A company produces two kinds of products. A product of the first type requires 1/4 hours of assembly labor, 1/8 hours of testing, and \$1.2 worth of raw materials. A product of the second type requires 1/3 hours of assembly, 1/3 hours of testing, and \$0.9 worth of raw materials. Given the current personnel of the company, there can be at most 90 hours of assembly labor and 80 hours of testing, each day. Products of the first and second type have a market value of \$9 and \$8, respectively.

(a) Formulate a linear programming problem that can be used to maximize the daily profit of the company.

(b) Consider the following two modifications to the original problem:

(i) Suppose that up to 50 hours of overtime assembly labor can be scheduled, at a cost of \$7 per hour.

(ii) Suppose that the raw material supplier provides a 10% discount if the daily bill is above \$300.

Which of the above two elements can be easily incorporated into the linear programming formulation and how? If one or both are not easy to incorporate, indicate how you might nevertheless solve the problem.

a) $9-1.2=7.8 \quad 8-0.9=7.1$

$$\begin{array}{ll} \text{Max} & 7.8x_1+7.1x_2 \\ \text{subject to} & \tfrac{1}{4}x_1+\tfrac{1}{3}x_2 \le 90 \\ & \tfrac{1}{8}x_1+\tfrac{1}{3}x_2 \le 80 \\ & x_1 \geq 0 \quad x_2 \geq 0 \end{array}$$

bi) $$\begin{array}{ll} \text{Max} & 7.8x_1+7.1x_2-7x_3 \\ \text{subject to} & \tfrac{1}{4}x_1+\tfrac{1}{3}x_2-x_3 \le 90 \\ & \tfrac{1}{8}x_1+\tfrac{1}{3}x_2 \le 80 \\ & x_1 \geq 0 \quad x_2 \geq 0 \\ & x_3 \le 50 \end{array}$$

ii) We could do two LPs with and without the discount to compare which would be better. One LP would have

$1.2x_1+0.9x_2 \le 300$

and the other with

$1.2x_1+0.9x_2 > 300$

with the modification of the maximizing function accordingly.

2. The function $f:\mathbb{R}^n \to \mathbb{R}$ is *convex* if for any $\boldsymbol{x}$, $\boldsymbol{y}$ in f's domain and $\lambda \in [0,1]$,

$$f(\lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{y}) \le \lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y}).$$

(a) (5 points) Prove that any local minimizer of f is a global minimizer.

(b) (5 points) f is *strictly convex* if the above inequality is strict. Show that if f has a minimizer on its domain, then that minimizer is unique.

Let x be the local min of f and $z \in \mathbb{R}^n$

Consider $y=\lambda z+(1-\lambda)x$ for $\lambda \in (0,1]$

Since $\mathbb{R}^n$ is convex, $y \in \mathbb{R}^n$.

It follows

$f(x) \le f(y) = f(\lambda z+(1-\lambda)x) \le \lambda f(z)+(1-\lambda)f(x)$

$\lambda f(x) \leq \lambda f(z)$

If $\lambda = 0$ $y = x$ which is a min

Since z is arbitrary x is the global min

b) Assume x and y are minimizers. I am assuming f is convex.

Let $z = \frac{x+y}{2}$

$f(x) \leq f(z) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$ $\qquad$ $f(y) \leq f(z) \leq \frac{1}{2}f(x) + \frac{1}{2}(y)$

$\frac{1}{2}f(x) \leq \frac{1}{2}f(y)$ $\qquad$ $\frac{1}{2}f(y) \leq \frac{1}{2}f(x)$

$\Rightarrow f(x) = f(y)$ $\therefore$ f has a unique min

3. The *Minkowski sum* of sets $S, T \subseteq \mathbb{R}^n$ is

$$V = \{\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} : \boldsymbol{x} \in S \text{ and } \boldsymbol{y} \in T\}.$$

(a) (5 points) Show that if S and T are convex, V is convex.

(b) (5 points) Does the converse hold? Prove or give a counterexample.

a) Let $\lambda \in [0,1]$ be fixed

Let $s = \lambda s_x + (1-\lambda)s_y \in S$ and $t = \lambda t_x + (1-\lambda)t_y \in T$

Consider $z = s + t = \lambda s_x + (1-\lambda)s_y + \lambda t_x + (1-\lambda)t_y$

$= \lambda(s_x + t_x) + (1-\lambda)(s_y + t_y)$

Since $s+t \in V$ and $s+t = \lambda(s_x+t_x) + (1-\lambda)(s_y+t_y)$,

V is convex. $\left((s_x + t_x,\ s_y + t_y) \in V^2\right)$

b) Let $V = [0,1]$

$V = (0,1) + \{0,1\}$

$\{0,1\}$ is not convex.

4. Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and $c \in \mathbb{R}$, the *level set* of f with respect to c is $L^=(c) = \{\boldsymbol{x} : f(\boldsymbol{x}) = c\}$. The *lower level set* of f with respect to c is $L^{\leq}(c) = \{\boldsymbol{x} : f(\boldsymbol{x}) \leq c\}$.

(a) (5 points) Prove that if f is convex, then $L^{\leq}(c)$ is convex.

(b) (5 points) Prove that $L^=(c)$ is convex if c is the minimum value of f.

a) Let $x, y \in L^{\leq}(c) \Rightarrow f(x) \leq c$ $f(y) \leq c$

It follows $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ for $\lambda \in [0,1]$

$\leq \lambda c + (1-\lambda)c$

Thus $f(\lambda x+(1-\lambda)y)\leq c \Rightarrow \overset{=c}{L^{\leq}(c)}$ is convex

b) Let c be min val of f which is convex

Let $x, y \in L^{=}(c)$

$c \leq f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) = c \quad \lambda \in [0,1]$

Thus $\lambda x + (1-\lambda) y \in L^{=}(c)$

5. (5 points) Prove the well known inequality between the arithmetic and geometric means of a set of positive numbers:

$$(x_1 + x_2 + \cdots + x_k)/k \geq (x_1 x_2 \cdots x_k)^{1/k}.$$

(Hint: Apply Jensen's inequality to $f(x) = -\log(x)$.)

log is concave so

$\log\left(\frac{\sum x_i}{k}\right) \geq \frac{\sum \log x_i}{k}$ Jensen's Inequality for concave

$\Rightarrow e^{\log\left(\frac{\sum x_i}{k}\right)} \geq e^{\frac{\sum \log x_i}{k}}$ since e is a bijection

$\frac{\sum x_i}{k} \geq \left(e^{\sum \log x_i}\right)^{1/k}$

$= \left(\prod e^{\log x_i}\right)^{1/k}$

$= \left(\prod x_i\right)^{1/k}$

6. (5 points) Let f be a twice continuously differentiable function. Recall the second-order Taylor expansion at $\boldsymbol{x}$ is

$$f(\boldsymbol{x}+\boldsymbol{d}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}) \boldsymbol{d} + o(\|\boldsymbol{d}\|^2).$$

Prove that f is convex if and only if the Hessian of f is positive semidefinite at any $\boldsymbol{x}$ in the domain of f.

($\Leftarrow$) Define $g(t) = f(x+td)$

$g''(t) = d^T \nabla^2 f d \geq 0$ By assumption

$\Rightarrow g'(t)$ is monotonic increasing $\forall t \in [a,b]$

$\Rightarrow \frac{g(b)-g(a)}{b-a} \geq g'(t)$

$g(b) \geq g(a) + g'(t)(b-a)$

$\therefore g$ is convex by prop in notes

$\Rightarrow f$ is convex

$(\Rightarrow)$ $f(x+d) = f(x) + \nabla f^T(x)\, d + \frac{1}{2} d^T \nabla^2 f(x) d + o(\|d\|^2)$

Let $y = x+d$ $\ni$

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

$\leq f(y) + \frac{1}{2}(y-x)^T \nabla^2 f(x)(y-x) + o(\|d\|^2)$ by prop since f is convex

$0 \leq \frac{1}{2} d^T \nabla^2 f(x) d + o(\|d\|^2)$

$0 \leq \lim_{d \to 0} \frac{1}{2} d^T \nabla^2 f(x) d + o(\|d\|^2) \Rightarrow 0 \leq d^T \nabla^2 f(x) d$

7. (5 points) Find the minimum of $f(\boldsymbol{x}) = \frac{1}{3}(x_1^2 + x_2^2)$ using the steepest descent algorithm with initial vector $\boldsymbol{x} = (1, 2)$ and constant $\lambda = 1$.

$\nabla f = \left(\frac{2}{3}x_1, \frac{2}{3}x_2\right) = \frac{2}{3}x$ $\quad \nabla^2 f(x) = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{2}{3} \end{bmatrix} \succ 0 \Rightarrow \sigma_p(\nabla^2 f(x)) = \{\frac{2}{3}\}$

$\Rightarrow$ f is convex since $\frac{2}{3} > 0$

Thus there is a unique global min

$x^* = x^* - \nabla f(x^*) = x^* - \frac{2}{3}x^*$

$x^* = \frac{1}{3}x^* \Rightarrow x^* = 0$

8. Consider the following unconstrained nonlinear program:

$$f(\boldsymbol{x}) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

(a) (5 points) What are the gradient and Hessian of f?

(b) (5 points) Starting from the initial point $\begin{bmatrix} 1 & 2 \end{bmatrix}^T$, find a stationary point $\boldsymbol{x}^*$ using the steepest descent method with constant step size $\alpha = 1/4$.

(c) (5 points) Is $\boldsymbol{x}^*$ an optimal solution? if so, is it unique? Explain.

a) $\nabla f(x) = Ax + b = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ -6 \end{bmatrix}$

$\nabla^2 f(x) = A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$

Note $\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \succ 0$ since $\operatorname{tr} A = 6$ $\det A = 8 \Rightarrow \sigma_p(A) = \{4, 2\}$

$4 > 0$ $2 > 0$

b)

```
x=zeros([20,2]);
x(1,:)=[1,2];
df=@(x) [3/4 -1/4;-1/4 3/4]*x'+[1/2;-3/2];

for k=1:19
```

As you can see

$x^* = (0, 2)$

b)
```
x=zeros([20,2]);
x(1,:)=[1,2];
df=@(x) [3/4 -1/4;-1/4 3/4]*x'+[1/2;-3/2];

for k=1:19
    x(k+1,:)=x(k,:)-df(x(k,:))';
end
```

```
x =

    1.0000    2.0000
    0.2500    2.2500
    0.1250    2.1250
    0.0625    2.0625
    0.0312    2.0312
    0.0156    2.0156
    0.0078    2.0078
    0.0039    2.0039
    0.0020    2.0020
    0.0010    2.0010
    0.0005    2.0005
    0.0002    2.0002
    0.0001    2.0001
    0.0001    2.0001
    0.0000    2.0000
    0.0000    2.0000
    0.0000    2.0000
    0.0000    2.0000
    0.0000    2.0000
    0.0000    2.0000
```

As you can see $x^* = (0,2)$

c) Since $A \succ 0$ this is a unique optimal solution

9. (5 points) Perform one Newton iteration for finding the minimum of $f(x,y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$ using initial vector $\boldsymbol{x} = (2,2)$.

$$x_{k+1} = x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k)$$

$$\nabla f(x) = (4x^3y^2 + 2xy^4 - 6xy^2,\ 2x^4y + 4x^2y^3 - 6x^2y)$$

$$\nabla^2 f(x) = \begin{bmatrix} 12x^2y^2 + 2y^4 - 6y^2 & 8x^3y + 8xy^3 - 12xy \\ 8x^3y + 8xy^3 - 12xy & 2x^4 + 12x^2y^2 - 6x^2 \end{bmatrix}$$

$$X_1 = X_0 - \nabla^2 f(X_0)^{-1} \nabla f(X_0) = \begin{bmatrix} \frac{28}{17} \\ \frac{28}{17} \end{bmatrix}$$

$$X_1 = X_0 - \nabla^2 f(X_0)\nabla f(X_0) = \begin{bmatrix} \frac{\quad}{17} \\ \frac{28}{17} \end{bmatrix}$$

```
f(x,y):=x^4·y^2+x^2·y^4−3·x^2·y^2+1;
df(a,b):=[at(at(diff(f(x,y),x),x=a),y=b),at(at(diff(f(x,y),y),x=a),y=b)];
H(a,b):=m: at(at(matrix ([diff(diff(f(x,y),x),x),diff(diff(f(x,y),x),y)],[diff(diff(f(x,y),x),y),diff(diff(f(x,y),y),y)]),x=a),y=b);
```

$f(x,y):=x^4 y^2 + x^2 y^4 + -3x^2 y^2 + 1$

$df(a,b):=\left[at\left(at\left(\frac{d}{dx} f(x,y), x=a\right), y=b\right), at\left(at\left(\frac{d}{dy} f(x,y), x=a\right), y=b\right)\right]$

$H(a,b):=m:at\left(at\left(\begin{pmatrix} \frac{d}{dx}\left(\frac{d}{dx} f(x,y)\right) & \frac{d}{dy}\left(\frac{d}{dx} f(x,y)\right) \\ \frac{d}{dy}\left(\frac{d}{dx} f(x,y)\right) & \frac{d}{dy}\left(\frac{d}{dy} f(x,y)\right) \end{pmatrix}, x=a\right), y=b\right)$

This is a language called Maxima

```
[[2],[2]]−invert(H(2,2)).df(2,2);
```

$$\begin{pmatrix} \frac{28}{17} \\ \frac{28}{17} \end{pmatrix}$$

10. (5 points) Let $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T A \boldsymbol{x} + \boldsymbol{b}^T \boldsymbol{x} + c$ for some positive definite A. Show that Newton's method to find a minimum of f starting from an arbitrary initial vector $\boldsymbol{x}^0$ in a single iteration. Give an intuitive explanation for this finding.

$$X_1 = X_0 - \nabla^2 f(X_0)^{-1} \nabla f(X_0)$$
$$= X_0 - A^{-1}(AX_0 + b)$$
$$= \cancel{X_0 - X_0}^{\,0} - A^{-1} b$$
$$X_1 = -A^{-1} b$$

Intuitively, Newton's method was constructed as a Taylor's expansion and thus will solve this perfectly