THE UNIVERSITY OF MELBOURNE

Student Number	

Semester 1 Assessment, 2020

School of Mathematics and Statistics

MAST30013 Techniques in Operations Research

This exam consists of 19 pages (including this page)

Authorised materials: printed one-sided copy of the Exam or the Masked Exam made available earlier, or an offline electronic pdf reader, up to two double-sided A4 pages of notes, school approved calculators, and blank A4 paper

Instructions to Students

- During exam writing time you may only interact with the device running the Zoom session with supervisor permission. The screen of any other device must be visible in Zoom from the start of the session.
- If you have a printer, print out the exam single-sided and hand write your solutions into the answer spaces.
- If you do not have a printer, or if your printer fails on the day of the exam,

  (a) download the exam paper to a second device (not running Zoom), disconnect it from the internet as soon as the paper is downloaded and read the paper on the second device;

  (b) write your answers on the Masked Exam PDF if you were able to print it single-sided before the exam day.

  If you do not have the Masked Exam PDF, write single-sided on blank sheets of paper.
- If you are unable to answer the whole question in the answer space provided then you can append additional handwritten solutions to the end of your exam submission. If you do this you MUST make a note in the correct answer space or page for the question, warning the marker that you have appended additional remarks at the end.
- Assemble all the exam pages (or template pages) in correct page number order and the correct way up, and add any extra pages with additional working at the end.
- Scan your exam submission to a single PDF file with a mobile phone or a scanner. Scan from directly above to avoid any excessive keystone effect. Check that all pages are clearly readable and cropped to the A4 borders of the original page. Poorly scanned submissions may be impossible to mark.
- Upload the PDF file via the Canvas Assignments menu and submit the PDF to the GradeScope tool by first selecting your PDF file and then clicking on Upload PDF.
- Confirm with your Zoom supervisor that you have GradeScope confirmation of submission before leaving Zoom supervision.
- You are not permitted to use electronic equipment other than for running Zoom and reviewing the exam questions.
- You are permitted to use up to 2 double-sided A4 pages of notes, which may be printed and/or handwritten.
- No formula sheet is provided with this exam paper.
- There are 6 questions with marks as shown. The total number of marks available is 80.

Question 1 (15 marks)

Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = x^4 + 2x^2 + 3x.$$

(a) Show that f is continuous and unimodal over its domain $\mathbb{R}$.

f is a polynomial of x and hence continuous over its domain $\mathbb{R}$.
$f'(x) = 4x^3 + 4x + 3$ and $f''(x) = 12x^2 + 4$.
$f''(x) > 0$ for $x \in \mathbb{R}$ and hence f is strictly convex and unimodal.

(b) The minimum of f is in the interval $(-\infty, 0]$ (you do not need to prove that). Find a number a, such that the minimum of f lies in $[a, 0]$.

Choose a step size $T = 1$. Then

$$f(0 - T) = f(-1) = 0 = f(0),$$

$$f(0 - 2T) = f(-2) = 18 > f(0).$$

So $a = -2$.

(c) Starting with the interval $[-2, 0]$ and using the Golden Section Search method, what is the uncertainty interval size after 6 f-calculations?

The final interval size is

$$(b - a)\gamma^{n-1} = 2 \times 0.618^5 \approx 0.180$$

Page 3 of 19 — add extra pages after page 19 — Page 3 of 19

(d) Perform one iteration of the Fibonacci Search method in finding the minimum of f over the interval $[-2, 0]$ to a tolerance of $\epsilon = 0.01$.

Fibonacci numbers

n	0	1	2	3	4	5	6	7	8	9	10	11	12	...
F_n	1	1	2	3	5	8	13	21	34	55	89	144	233	...

$$\frac{b-a}{F_n} = \frac{2}{F_n} < 2\epsilon = 0.02$$
$$\Rightarrow F_n > 100 \Rightarrow n = 11$$
$$p = b - \frac{F_{10}}{F_{11}}(b-a) = 0 - \frac{89}{144} \times 2 \approx -1.236$$
$$q = a + \frac{F_{10}}{F_{11}}(b-a) = -2 + \frac{89}{144} \times 2 \approx -0.764$$
$$f(p) \approx 1.681 > f(q) \approx -0.784$$
$$\Rightarrow a = -1.236$$

The uncertainty interval becomes $[-1.236, 0]$ after 1 iteration of Fibonacci search.

(e) Can the False Position method be used to find the minimum of f over the interval $[-2, 0]$? Explain why or why not.

$f'(x)$ is an increasing function on $[-2, 0]$ since $f''(x) > 0$.
$f'(0) = 3 > 0$ and $f'(-2) = -37 < 0$. So we can apply the false position method.

(f) Can we apply Newton's method in finding the minimum of f over the interval $[-2, 0]$, starting from $x_0 = -1$? If so, perform one iteration. Otherwise, explain why not.

Yes, because $f''(x) = 12x^2 + 4$ is not small for all $x \in \mathbb{R}$.

$$\begin{aligned} a &= x_0 = -1 \\ g(a) &= f'(a) = -5 \\ g'(a) &= f''(a) = 16 \\ p &= a - \frac{g(a)}{g'(a)} = -1 - \frac{-5}{16} = -\frac{11}{16} \end{aligned}$$

The new estimate is $-\frac{11}{16}$.

Question 2 (13 marks)

Consider the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x) = x_1^2 x_2 - 2x_1x_2 - x_2^2 - 3x_2 + 6.$$

(a) Using the first-order necessary condition, find all stationary points of f.

$$\nabla f(x) = \begin{bmatrix} 2x_1x_2 - 2x_2 \\ x_1^2 - 2x_1 - 2x_2 - 3 \end{bmatrix} = 0$$

From the 1st equation, $x_2 = 0$ or $x_1 = 1$.
Substituting $x_2 = 0$ into the 2nd equation yields $x_1 = 3$ or -1.
Substituting $x_1 = 1$ into the 2nd equation yields $x_2 = -2$.

There are 3 stationary points: $(3, 0)$, $(-1, 0)$ and $(1, -2)$.

(b) Using the second-order sufficiency condition, determine whether the stationary points found in (a) are local minimums, local maximums, or saddles points.

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

At $x^* = (3, 0)^T$,

$$\nabla^2 f(x^*) = \begin{bmatrix} 0 & 4 \\ 4 & -2 \end{bmatrix}$$

$$\begin{vmatrix} -\lambda & 4 \\ 4 & -2-\lambda \end{vmatrix} = 0$$

$$\lambda(\lambda + 2) - 16 = 0 \Rightarrow \lambda = -1 \pm \sqrt{17}$$

There is extra answer space for Question 2 on the next page

Additional answer space only for Question 2—submit this page even if not used

Given $\nabla^2 f(x^*)$ has 1 positive eigenvalue and 1 negative eigenvalue (and hence is neither positive definite or negative definite), $(3,0)$ is a saddle point.
At $x^* = (-1,0)^T$,

$$\nabla^2 f(x^*) = \begin{bmatrix} 0 & -4 \\ -4 & -2 \end{bmatrix}$$

Then

$$\lambda(\lambda+2) - 16 = 0 \Rightarrow \lambda = -1 \pm \sqrt{17}$$

Given $\nabla^2 f(x^*)$ has 1 positive eigenvalue and 1 negative eigenvalue, $(-1,0)$ is a saddle point.

At $x^* = (1,-2)^T$,

$$\nabla^2 f(x^*) = \begin{bmatrix} -4 & 0 \\ 0 & -2 \end{bmatrix}$$

is a diagonal matrix with negative diagonal components (or has $\lambda_1 = -4$ and $\lambda_2 = -2$) and thus is negative definite.
$(1,-2)$ is a local maximum.

(c) Write down the second-order Taylor approximation around a saddle point you found in (b) (if more than one saddle point were found, just consider one of them). Using the second-order Taylor approximation, verify that the value of f decreases along direction $d = (0, 1)^T$.

Hint: The second-order Taylor approximation around x^0 is

$$f(x) \approx f(x^0) + \nabla f(x^0)(x - x^0) + \frac{1}{2}(x - x^0)^T \nabla^2 f(x^0)(x - x^0).$$

The second-order Taylor approximation at a stationary point is

$$f(x) \approx f(x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*)$$

At $x^* = (3, 0)^T$, for some x near x^*

$$f(x) \approx 6 + \frac{1}{2}(x - (3, 0))^T \begin{bmatrix} 0 & 4 \\ 4 & -2 \end{bmatrix} (x - (3, 0))$$

Given $x = x^* + \epsilon(0, 1)^T$ for some $\epsilon > 0$,

$$f(x) \approx 6 + \frac{1}{2}\epsilon(0, 1)^T \begin{bmatrix} 0 & 4 \\ 4 & -2 \end{bmatrix} \epsilon(0, 1) = 6 - \epsilon^2 < 6 = f(x^*)$$

Hence $d = (0, 1)^T$ is a descent direction at the saddle point.
Alternatively, at $x^* = (-1, 0)^T$, for some x near x^*

$$f(x) \approx 6 + \frac{1}{2}(x - (-1, 0))^T \begin{bmatrix} 0 & -4 \\ -4 & -2 \end{bmatrix} (x - (-1, 0))$$

Given $x = x^* + \epsilon(0, 1)^T$ for some $\epsilon > 0$,

$$f(x) \approx 6 + \frac{1}{2}\epsilon(0, 1)^T \begin{bmatrix} 0 & -4 \\ -4 & -2 \end{bmatrix} \epsilon(0, 1) = 6 - \epsilon^2 < 6 = f(x^*)$$

Question 3 (15 marks)

Consider the unconstrained nonlinear program

$$\min_x f(x) = \frac{1}{2}(x_1+2)^2 + \frac{1}{2}(x_2-1)^2 + x_3^2.$$

(a) Perform one iteration of the Steepest Descent method starting at the point $x^0 = (-1, 0, 1)^T$ to find x^1. Choose the step size by optimising the single-variable function. What is the angle between $-\nabla f(x^0)$ and $-\nabla f(x^1)$?

$$\nabla f(x) = \begin{bmatrix} x_1+2 \\ x_2-1 \\ 2x_3 \end{bmatrix}$$

$$\nabla f(x^0) = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

$$d = -\nabla f(x^0) = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$$

$$x^1 = x^0 + dt = \begin{bmatrix} -1-t \\ t \\ 1-2t \end{bmatrix}$$

$$\begin{aligned} f(t) = f(x^1) &= \frac{1}{2}(1-t)^2 + \frac{1}{2}(t-1)^2 + (1-2t)^2 \\ &= 5t^2 - 6t + 2 \\ f'(t) &= 10t - 6 = 0 \Rightarrow t^* = \frac{3}{5} \end{aligned}$$

So $x^1 = (-\frac{8}{5}, \frac{3}{5}, -\frac{1}{5})^T$.

The angle between $-\nabla f(x^0)$ and $-\nabla f(x^1)$ is $\pi/2$ since in the steepest descent method, $d^k = -\nabla f(x^k)$ is normal to $d^{k+1} = -\nabla f(x^{k+1})$.

(b) Show that the step size found in part (a) satisfies the Armijo-Goldenstein condition with $\sigma = \frac{1}{3}$.

The Armijo-Goldenstein condition states $f(t) \leq f(0) + t\sigma f'(0)$.

$$\begin{aligned} f(0) + t^*\sigma f'(0) &= 2 + \frac{3}{5} \times \frac{1}{3} \times (-6) = \frac{4}{5} \\ &\geq f(t^*) = 5 \times \left(\frac{3}{5}\right)^2 - 6 \times \frac{3}{5} + 2 = \frac{1}{5} \end{aligned}$$

Hence the condition is satisfied.

(c) Find the Newton direction at the point $x^0 = (-1, 0, 1)^T$. Is the Newton direction a descent direction of f at x^0? Justify your answer.

$$\nabla^2 f(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$\nabla^2 f(x)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$$

$$\begin{aligned} d &= -\nabla^2 f(x)^{-1} \nabla f(x^0) \\ &= \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

Evaluating the directional derivative

$$\langle \nabla f(x^0), d \rangle = [1 \quad -1 \quad 2] \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} = -4 < 0$$

There is extra answer space for Question 3 on the next page

Additional answer space only for Question 3—submit this page even if not used

Thus, the Newton direction is a descent direction.

Alternatively,

(i) because $\nabla f(x^0) \neq 0$ and $\nabla^2 f(x^0)$ is positive definite as it is a diagonal matrix with positive diagonal components, the Newton direction is a descent direction by the lemma of conditions for the Newton direction.

(ii) Because $\nabla^2 f(x^0)$ is positive definite as it is a diagonal matrix with positive diagonal components, $\nabla^2 f(x^0)^{-1}$ exists and is also positive definite. Then

$$\begin{aligned}\langle \nabla f(x^0), d\rangle &= \langle \nabla f(x^0), -\nabla^2 f(x^0)^{-1}\nabla f(x^0)\rangle \\ &= -\nabla f(x^0)^T \nabla^2 f(x^0)^{-1} \nabla f(x^0) \\ &< 0\end{aligned}$$

by positive definiteness of $\nabla^2 f(x^0)^{-1}$.

(d) Explain why only one iteration of the Newton's method would be required to find the global minimum of f.

Because f is a quadratic function with positive definite Hessian.

(e) Find the BFGS direction for f at the point $x^0 = (-1, 0, 1)^T$ with $H_0 = I_3$, where I_3 is the 3×3 identity matrix.

The BFGS direction is the same as the steepest descent direction $d = (-1, 1, -2)^T$.

Question 4 (16 marks)

Consider the constrained nonlinear program

$$\begin{array}{rc} \min_x & f(x) = (x_1 - 2)^2 + (x_2 - 2)^2 \\ s.t. & x_1 x_2 \leq -3 \\ & 2x_1 - x_2 \leq -5 \end{array}$$

(a) Write down the Lagrangian of the program.

$g_1(x) := x_1 x_2 + 3 \leq 0$ and $g_2(x) := 2x_1 - x_2 + 5 \leq 0$.

$$L(x, \lambda) = (x_1 - 2)^2 + (x_2 - 2)^2 + \lambda_1(x_1 x_2 + 3) + \lambda_2(2x_1 - x_2 + 5)$$

(b) Show that $x^* = (-1, 3)^T$ is a stationary point of the program. Find the associated optimal KKT multiplier λ^*.

KKTa:

$$\nabla_x L(x, \lambda) = \begin{bmatrix} 2(x_1 - 2) + \lambda_1 x_2 + 2\lambda_2 \\ 2(x_2 - 2) + \lambda_1 x_1 - \lambda_2 \end{bmatrix} = 0 \qquad \begin{array}{c} (1) \\ (2) \end{array}$$

KKTb:

$$\lambda_1, \lambda_2 \geq 0 \tag{3}$$
$$x_1 x_2 + 3 \leq 0 \tag{4}$$
$$2x_1 - x_2 + 5 \leq 0 \tag{5}$$
$$\lambda_1(x_1 x_2 + 3) = 0 \tag{6}$$
$$\lambda_2(2x_1 - x_2 + 5) = 0 \tag{7}$$

KKTc: no equality constraint

There is extra answer space for Question 4 on the next page

Additional answer space only for Question 4—submit this page even if not used

At x^*, (4) and (5) obtain equality and furthermore (6) and (7) hold.

$$\nabla_x L(x^*, \lambda) = \begin{bmatrix} -6 + 3\lambda_1 + 2\lambda_2 \\ 2 - \lambda_1 - \lambda_2 \end{bmatrix} = 0$$

Solving the above gives $\lambda_1^* = 2 > 0$ and $\lambda_2^* = 0 \geq 0$.
Then (1) - (3) are satisfied.
So, x^* is a stationary point and further the KKT point is $(x^*, \lambda^*) = ((-1, 3), (2, 0))$.

(c) Check whether a constraint qualification holds at x^*.

Both constraints are active. $g_1(x)$ is not affine.

$$\nabla g_1(x) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \qquad \nabla g_2(x) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\nabla g(x^*) = \begin{bmatrix} 3 & 2 \\ -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{2}{3} \\ 0 & 1 \end{bmatrix}$$

has full rank. So LICQ holds.
Alternatively, for $d = (0,1)^T$, $(3,-1)d < 0$ and $(2,-1)d < 0$. So MFCQ holds.

(d) Find the critical cone of the KKT point (x^*, λ^*).

g_1 is active with $\lambda_1^* > 0$ and g_2 is active with $\lambda_2^* = 0$.
The critical cone at the KKT point is

$$\begin{aligned} \mathcal{C}(x^*, \lambda^*) &= \{d \in \mathbb{R}^2 : \nabla g_1(x^*)^T d = 0; \nabla g_2(x^*)^T d \leq 0\} \\ &= \{d \in \mathbb{R}^2 : d = (d_1, d_2), 3d_1 - d_2 = 0; 2d_1 - d_2 \leq 0\} \\ &= \{d \in \mathbb{R}^2 : d = (d_1, 3d_1),\ d_1 > 0\} \end{aligned}$$

(e) Check the second-order sufficient condition at the KKT point (x^*, λ^*) and determine if x^* is a local minimum.

$$\nabla^2_{xx} L(x,\lambda) = \begin{bmatrix} 2 & \lambda_1 \\ \lambda_1 & 2 \end{bmatrix} \qquad \nabla^2_{xx} L(x^*,\lambda^*) = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

For $d \in \mathcal{C}(x^*, \lambda^*)$

$$\begin{aligned} d^T \nabla^2_{xx} L(x^*,\lambda^*) d &= [d_1 \quad 3d_1] \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} d_1 \\ 3d_1 \end{bmatrix} \\ &= 32 d_1^2 > 0 \ \forall d_1 > 0 \end{aligned}$$

Given the Hessian is positive definite on the critical cone at the KKT point (x^*, λ^*), it is a local minimum.

Question 5 (12 marks)

Consider the constrained nonlinear program

$$\begin{aligned} \min_{x} \quad & f(x) = x_1^2 - x_2 \\ s.t. \quad & x_1 \geq 1 \\ & x_1 + x_2 = 1 \end{aligned}$$

(a) Write down the l_2 penalty function $P_k(x)$ with penalty parameter $\alpha_k = k$.

$$P_k(x) = x_1^2 - x_2 + \frac{k}{2}((1 - x_1)_+)^2 + \frac{k}{2}(x_1 + x_2 - 1)^2$$

(b) Write down $\nabla P_k(x)$ and solve $\nabla P_k(x) = 0$ to find the single stationary point $x^k = (x_1^k, x_2^k)$ of $P_k(x)$.

$$\nabla P_k(x) = \begin{bmatrix} 2x_1 - k(1 - x_1)_+ + k(x_1 + x_2 - 1) \\ -1 + k(x_1 + x_2 - 1) \end{bmatrix} = 0 \qquad \begin{matrix} (1) \\ (2) \end{matrix}$$

If $x_1 \geq 1$, then (1) becomes $2x_1 + 1 = 0$ by (2). Furthermore, $x_1 = -\frac{1}{2} < 1$. So, $x_1 < 1$.
Given $x_1 < 1$, (1) becomes $2x_1 - k(1 - x_1) + 1 = 0$. Then

$$\begin{aligned} x_1^k &= \frac{k-1}{k+2} \\ x_2^k &= \frac{1}{k} + 1 - \frac{k-1}{k+2} \\ &= \frac{k + 2 + k^2 + 2k - k^2 + k}{k(k+2)} \\ &= \frac{2(2k+1)}{k(k+2)} \end{aligned}$$

(c) Solve the program by finding the limit $x^* = \lim_{k\to\infty}(x_1^k, x_2^k)$.

$$
\begin{aligned}
x^* &= \lim_{k\to\infty} \left(\frac{k-1}{k+2}, \frac{2(2k+1)}{k(k+2)} \right) \\
&= (1, 0)
\end{aligned}
$$

(d) Write down an estimate (λ^k, η^k) of the optimal Lagrange multiplier vector. Find the limit $(\lambda^*, \eta^*) = \lim_{k\to\infty}(\lambda^k, \eta^k)$.

$$
\lambda^k = k(1 - x_1^k)_+, \qquad \eta^k = k(x_1^k + x_2^k - 1)
$$

Since $k(x_1^k + x_2^k - 1) = 1$ by (2),

$$
\begin{aligned}
(\lambda^*, \eta^*) &= \lim_{k\to\infty} (k(1 - x_1^k)_+, k(x_1^k + x_2^k - 1)) \\
&= \lim_{k\to\infty} \left(k \left(1 - \frac{k-1}{k+2} \right)_+, 1 \right) \\
&= \lim_{k\to\infty} \left(\frac{3k}{k+2}, 1 \right) \\
&= (3, 1)
\end{aligned}
$$

(e) What difficulty arises if we attempt to employ second-order search methods, such as Newton's method, in order to estimate the minimum of the penalty function $P_k(x)$?

Second-order search methods require the function to be minimised to be C^2. As $P_k(x)$ is C^1 but not C^2, we cannot apply second-order search methods to solve $\min P_k(x)$.

Question 6 (9 marks)

Consider the following convex program

$$\begin{aligned} \min_x \quad & f(x) \\ s.t. \quad & g_i(x) \leq 0, \quad \text{for } i = 1, \ldots, p. \end{aligned}$$

Let $L(x, \lambda)$ be the Lagrangian for the program and let (x^*, λ^*) be a KKT point of the program.

(a) Explain why $L(x, \lambda^*)$ is a convex function of x.

The program is convex and so $f(x)$ and $g_i(x)$ are all convex.
Given (x^*, λ^*) is a KKT point of the program, by KKTb, $\lambda_i^* \geq 0$, for $i = 1, \ldots, p$.
$L(x, \lambda^*) = f(x) + \sum_{i=1}^p \lambda_i^* g_i(x)$ is then convex.

(b) Show x^* minimises $L(x, \lambda^*)$.

Given (x^*, λ^*) is a KKT point of the program, by KKTa,

$$\nabla_x L(x^*, \lambda^*) = 0.$$

So x^* is a stationary point of $\min_x L(x, \lambda^*)$.
Given $L(x, \lambda^*)$ is convex, $\min_x L(x, \lambda^*)$ is a (unconstrained) convex program.
The stationary point of an unconstrained convex program is a local/global minimum of the program.

(c) Write down the Lagrangian dual of this program.

$$\begin{aligned} \max_{\lambda} \quad & \psi(\lambda) \\ s.t. \quad & \lambda_i \geq 0 \ \text{ for } i = 1, \ldots, p \end{aligned}$$

where $\psi(\lambda) = \min_x L(x, \lambda)(= \min_x(f(x) + \sum_{i=1}^{p} \lambda_i g_i(x)))$.

End of Exam—Total Available Marks = 80