Imperial College London

MATH97405 MATH70005

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May-June 2022

This paper is also taken for the relevant examination for the
Associateship of the Royal College of Science

Optimisation

Date: 11 May 2022

Time: 09:00 – 11:30 (BST)

Time Allowed: 2:30 hours

Upload Time Allowed: 30 minutes

This paper has 5 Questions.

Candidates should start their solutions to each question on a new sheet of paper.

Each sheet of paper should have your CID, Question Number and Page Number on the top.

Only use 1 side of the paper.

Allow margins for marking.

Any required additional material(s) will be provided.

Credit will be given for all questions attempted.

Each question carries equal weight.

SUBMIT YOUR ANSWERS AS SEPARATE PDFs TO THE RELEVANT DROPBOXES ON BLACKBOARD (ONE FOR EACH QUESTION) WITH COMPLETED COVERSHEETS WITH YOUR CID NUMBER, QUESTION NUMBERS ANSWERED AND PAGE NUMBERS PER QUESTION.

1. (a) Let $\{x_1, \ldots, x_n\}$ be a random sample where all $x_i > 0$. We define the function from $f : \mathbb{R}_{++} \to \mathbb{R}$, where $\mathbb{R}_{++}$ denotes the positive orthant, as

$$f(\theta) = \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i}.$$

(i) Find a stationary point of f. (6 marks)
Hint: first treat this as an unconstrained optimization problem, then use that $\theta \in \mathbb{R}_{++}$.

(ii) Classify the point found in i): maximum, minimum, or saddle point only. (4 marks)

(b) Consider the problem

$$\min_{\mathbf{x} \in \mathbb{R}^2} \; \mathbf{x}^\top \begin{bmatrix} a & b \\ b & a \end{bmatrix} \mathbf{x}, \qquad a, b \in \mathbb{R}.$$

(i) Determine conditions on a and b for this problem to have a unique stationary point and find it. (4 marks)

(ii) Determine conditions on a and b such that the unique stationary point found in part i) is: a global minimizer, a global maximizer, or a saddle point. (6 marks)

(Total: 20 marks)

2. (a) Given a matrix $\mathbf{Q} \succ 0$ in $\mathbb{R}^{n \times n}$, let $\mathbf{x}^* \in \mathbb{R}^n$ be solution of

$$\max_{\mathbf{x}} \{\mathbf{x}^\top \mathbf{Q} \mathbf{x}, \text{ subject to } \|\mathbf{x}\|_\infty \leq 1\}$$

(i) How do we know there exists such an $\mathbf{x}^*$? (6 marks)

(ii) Consider the function $f : \mathbb{R}^n \to \mathbb{R}$ defined as

$$f(\mathbf{x}) = \prod_{i=1}^{n} |x_i|.$$

Determine the value of $f(\mathbf{x}^*)$. (4 marks)

(b) Determine whether the the following functions are coercive:

(i) $f(x_1, x_2) = 2x_1^2 - 8x_1x_2 + x_2^2$. (5 marks)

(ii) $h_\gamma(\mathbf{x}) = g(\mathbf{x}) + \gamma(\mathbf{x}^\top \mathbf{x})$, where $g : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and convex and $\gamma > 0$. (5 marks)
Hint: use the first-order characterization of convexity and Cauchy-Schwarz.

(Total: 20 marks)

3. (a) Determine whether the following functions are convex over $\mathbb{R}^n$:

(i) The function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ defined by $f(\mathbf{x}) = \exp(\mathrm{Tr}(\mathbf{x}\mathbf{x}^\top))$. (4 marks)

(ii) The function $g(\mathbf{x})$ such that

$$g(\mathbf{x}) = \begin{cases} 0 & \text{for } \mathbf{x} \in K \\ \|\mathbf{x}\| - 1 & \text{elsewhere} \end{cases},$$

where $K := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq 1\}$. (4 marks)

(iii) $f(\mathbf{x}) + g(\mathbf{x})$ where $f(\mathbf{x})$ and $g(\mathbf{x})$ are as in i) and ii) above, over $\mathbb{R}^n$. (2 marks)

(b) Consider the minimization problem

$$\min \left\{ f(\mathbf{x}) : \mathbf{a}^\top \mathbf{x} = 1, \mathbf{x} \in \mathbb{R}^n \right\} \tag{P}$$

where f is a continuously differentiable function over $\mathbb{R}^n$ and $\mathbf{a} \in \mathbb{R}^n_{++}$. Show that $\mathbf{x}^*$ satisfying $\mathbf{a}^\top \mathbf{x}^* = 1$ is a stationary point of (P) if and only if (5 marks for each direction)

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{a_1} = \frac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{a_2} = \cdots = \frac{\frac{\partial f}{\partial x_n}(\mathbf{x}^*)}{a_n}$$

(10 marks)

(Total: 20 marks)

4. (a) Consider the problem

$$\min\ x_1^2 + 2x_2^2 + x_1$$
$$\text{subject to } \ x_1 + x_2 \le \gamma\,,$$

(i) Show that for any $\gamma \in \mathbb{R}$, this problem has a unique solution. (4 marks)

(ii) Solve the problem using KKT conditions, expressing the general solution as a function of γ. (6 marks)

(b) Given a vector $\mathbf{y} \in \mathbb{R}^n$ and $\lambda > 0$, consider the sparse regularisation problem

$$\min_{\mathbf{x},\mathbf{w}}\ \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{w}\|_1$$
$$\text{subject to } \mathbf{w} = \mathbf{L}\mathbf{x}\,,$$

where where $\mathbf{L} \in \mathbb{R}^{(n-1)\times n}$ is given by

$$\mathbf{L} = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}.$$

(i) Using duality arguments and that

$$\min_{\mathbf{w}} \lambda\|\mathbf{w}\|_1 - \mu^\top \mathbf{w} = \begin{cases} 0\,, & \|\mu\|_\infty \le \lambda\,, \\ -\infty & \text{else}\,, \end{cases}$$

show that the dual problem is given by

$$\max_{\mu}\ -\frac{1}{4}\mu^\top \mathbf{L}\mathbf{L}^\top \mu + \mu^\top \mathbf{L}\mathbf{y}$$
$$\text{subject to } \|\mu\|_\infty \le \lambda\,.$$

(6 marks)

(ii) Explain why in this case the dual formulation is useful to find a solution to the primal problem and explain how would you proceed. How are the solutions of the primal and dual problems related? (4 marks)

(Total: 20 marks)

5. (a) Find the optimal control signal $u(t)$ solving the problem

$$\begin{aligned} \underset{u(\cdot)}{\text{minimize}} \quad & \int_0^T x_2(t) + u^2(t)\, dt \\ \text{subject to} \quad & \dot{x}_1(t) = -x_1(t) + u(t) \\ & \dot{x}_2(t) = x_1(t) \\ & x_1(0) = a, x_2(0) = 0 \end{aligned}$$

for a fixed $T > 0$, and arbitrary a, using Pontryagin's Maximum Principle. (10 marks)

(b) Solve the following optimal control problem

$$\begin{aligned} \underset{u(\cdot)}{\text{minimize}} \quad & \frac{1}{2}\int_0^\infty u^2(t) + x^2(t) + 2x^4(t)\, dt \\ \text{subject to} \quad & \dot{x}(t) = x^3(t) + u(t) \\ & x(0) = x_0\,, \end{aligned}$$

using dynamic programming and the Hamilton-Jacobi-Bellman assuming the value function can be expressed as $V(x) = \alpha x^2 + \beta x^4$. (10 marks)

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2022

This paper is also taken for the relevant examination for the Associateship.

MATH60005/70005/97405

Optimisation (Solutions)

Setter's signature	Checker's signature	Editor's signature
....	....	....

1. (a) (i) We first not that stationarity is equivalent to $f'(\theta) = 0$. Computing the derivative for a generic term inside $f(\theta)$ meth seen $\Downarrow$

$$\frac{d}{d\theta}\left(\frac{\theta^{x_i}e^{-\theta}}{x_i}\right) = \frac{\theta^{x_i}e^{-\theta}}{x_i}\left(\frac{x_i}{\theta} - 1\right).$$

leads to

$$f'(\theta) = \sum_{i=1}^{n}\left(\frac{x_i}{\theta} - 1\right) f(\theta).$$

However, we note that $f(\theta) > 0$ for all $\theta \in \mathbb{R}_{++}$, hence

$$f'(\theta) = 0 \Leftrightarrow \sum_{i=1}^{n}\left(\frac{x_i}{\theta} - 1\right) = 0 \Rightarrow \theta^* = \frac{1}{n}\sum_{i=1}^{n} x_i\,,\ \theta^* \in \mathbb{R}_{++}\,.$$

6, A

(ii) We need to check the sign of $f''(\theta^*)$:

$$f''(\theta^*) = f'(\theta^*)\sum_{i=1}^{n}\left(\frac{x_i}{\theta^*} - 1\right) - f(\theta^*)\sum_{i=1}^{n}\frac{x_i}{\theta^{*2}} = -f(\theta^*)\sum_{i=1}^{n}\frac{x_i}{\theta^{*2}} < 0\,,$$

as all the terms in the last expression are positive, hence θ^* is a maximum (it is in fact the maximum likelihood estimator of a Poisson distribution). 4, A

(b) (i) For a unique fixed point we need the matrix seen $\Downarrow$

$$Q := \begin{bmatrix} a & b \\ b & a \end{bmatrix}$$

to be invertible. This is equivalent to requiring that $\det(Q) = a^2 - b^2 \neq 0$, meaning $|a| \neq |b|$. In this case it is clear that the stationary point is $(0, 0)$. 4, A

(ii) Assuming Q invertible, a unique global maximizer is equivalent to $Q \succ 0$. For a 2-by-2 matrix this can be stated in terms of $\det(Q) = a^2 - b^2 > 0$ and $\mathrm{Tr}(Q) = 2a > 0$, or $a > 0$ and $|a| > |b|$. A global minimizer is characterized similarly, with $Q \prec 0$, meaning $a < 0$ and $|a| > |b|$. Assuming Q invertible, the only possibility for a saddle point is that the eigenvalues of Q have different sign, expressed as $\det(Q) = a^2 - b^2 < 0$, or $|a| < |b|$. 6, B

2. (a) (i) Since $\mathbf{Q} \succ 0$ it follows that the objective function is convex. The constraint set is convex, it's ℓ_∞ unit ball, and compact, so it follows there exists a maximizer at an extreme point of the constraint set. $\boxed{\text{sim. seen} \Downarrow}$ $\boxed{6, \text{A}}$

(ii) The exteme points of the ℓ_∞ unit ball are points in $\mathbb{R}^n$ such that each coordinate is in $\{-1, 1\}$, and hence $f(\mathbf{x}^*) = 1$ for a maximizer. $\boxed{4, \text{C}}$

(b) (i) The function is a quadratic function $\boxed{\text{unseen} \Downarrow}$

$$f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{x}^\top \begin{bmatrix} 2 & -4 \\ -4 & 1 \end{bmatrix} \mathbf{x}\,,$$

hence f is coercive if and only if $\mathbf{A} \succ 0$, which it clearly isn't as $\mathrm{Tr}(\mathbf{A}) = 3$ and $\det(\mathbf{A}) = -14$. $\boxed{4, \text{A}}$

(ii) Coercive. Using the first-order characterization of convexity for g we have that

$$g(0) = \nabla g(0)^\top \mathbf{x} \leq g(\mathbf{x})\,,$$

and using Cauchy-Schwarz as indicated

$$\begin{aligned} -\|\nabla g(0)\|\|\mathbf{x}\| &\leq \nabla g(0)^\top \mathbf{x} \\ g(0) + \gamma\|\mathbf{x}\|^2 - \|\nabla g(0)\|\|\mathbf{x}\| &\leq \nabla g(0)^\top \mathbf{x} + g(0) + \gamma\|\mathbf{x}\|^2 \leq h_\gamma(\mathbf{x})\,. \end{aligned}$$

From here, it is clear that when $\|\mathbf{x}\| \to \infty$, the quadratic term on the left overtakes $\|\mathbf{x}\|$, and hence $\lim_{\|\mathbf{x}\|\to\infty} h_\gamma(\mathbf{x}) = \infty$. $\boxed{6, \text{D}}$

3. (a) (i) Given a vector $\mathbf{x}$, $\mathrm{Tr}(\mathbf{x}\mathbf{x}^\top) = x_1^2 + x_2^2 + \ldots + x_n^2 = \|\mathbf{x}\|^2$, which is a convex function in $\mathbb{R}^n$. Composition with $\exp$ preserves convexity. seen ⇓ 4, A

(ii) For $g(\mathbf{x})$, we observe it corresponds to the distance function to the set K, which is a convex set (unit ball). The distance function to a convex set is convex (preservation of convexity under partial minimization). 4, B

(iii) The sum of two convex functions is convex. 2, A

(b) Given $\mathbf{a} \in \mathbb{R}^n_{++}$, we consider the feasible set $U = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^\top\mathbf{x} = 1\}$. A point $\mathbf{x}^* \in U$ is a stationary point if and only if meth seen ⇓

$$\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq 0\,, \quad \text{for all } \mathbf{x} \in U\,, \tag{I}$$

We will show that condition this is equivalent to the following simple and explicit condition

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{a_1} = \frac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{a_2} = \cdots = \frac{\frac{\partial f}{\partial x_n}(\mathbf{x}^*)}{a_n} \tag{II}$$

We begin by showing that (II) implies (I). Assume $\mathbf{x}^* \in U$ satisfies (II). Then, for any $\mathbf{x} \in U$ we have

$$\begin{aligned}\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}^*)\,(x_i - x_i^*) \\ &= \frac{\partial f}{\partial x_1}(\mathbf{x}^*)\frac{1}{a_1}\sum_{i=1}^n a_i(x_i - x_i^*) = \frac{\partial f}{\partial x_1}(\mathbf{x}^*)(\mathbf{a}^\top\mathbf{x} - \mathbf{a}^\top\mathbf{x}^*) = 0\,,\end{aligned}$$

and in particular $\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq 0$. We have thus shown that (I) is satisfied. 5, D

To show the reverse direction, take $\mathbf{x}^* \in U$ that satisfies (I) and assume in contradiction that (II) does not hold. This means that there exist two different indices $i \neq j$ such that

$$a_j \frac{\partial f}{\partial x_i}(\mathbf{x}^*) > a_i \frac{\partial f}{\partial x_j}(\mathbf{x}^*)\,.$$

Define the vector $\mathbf{x} \in U$ as

$$x_k = \begin{cases} x_k^*, & k \notin \{i,j\} \\ x_i^* - a_j, & k = i \\ x_j^* + a_i, & k = j \end{cases}\,.$$

There other alternatives for a vector, as long as the proposed alternative belongs to U. Then,

$$\begin{aligned}\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) &= \frac{\partial f}{\partial x_i}(\mathbf{x}^*)(x_i - x_i^*) + \frac{\partial f}{\partial x_j}(\mathbf{x}^*)\left(x_j - x_j^*\right) \\ &= -a_j \frac{\partial f}{\partial x_i}(\mathbf{x}^*) + a_i \frac{\partial f}{\partial x_j}(\mathbf{x}^*) \\ &< 0\end{aligned}$$

which is a contradiction to the assumption that (I) is satisfied. We thus conclude that (I) implies (II). 5, D

4. (a) (i) The cost is equivalent to minimizing a quadratic function $\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}$, with

unseen ⇓

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = [1\ 0]^\top .$$

It is clear that the eigenvalues of $\mathbf{A}$ are 1 and 2, and therefore this is a strictly convex function, with an affine constraint for any γ (which is a closed set). The cost is coercive. Under these conditions, there exists a unique optimal solution.

4, A

(ii) We proceed via KKT conditions, which in this case are necessary and sufficient. The Lagrangian is given by

$$L(\mathbf{x}, \lambda) = x_1^2 + 2x_2^2 + x_1 + \lambda(x_1 + x_2 - \gamma),$$

leading to $(\nabla_{\mathbf{x}} L = 0)$

$$\begin{aligned} 2x_1 + 1 + \lambda &= 0, \\ 4x_2 + \lambda &= 0, \\ \lambda(x_1 + x_2 - \gamma) &= 0, \\ \lambda &\geq 0. \end{aligned}$$

We analyse this system by cases. If $\lambda = 0$, then it follows that $x_1 = -1/2$ and $x_2 = 0$. This is a feasible solution if and only if $\gamma \geq -1/2$. In the second case, we assume $\lambda > 0$, which leads to the system

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma \\ -1 \\ 0 \end{bmatrix}.$$

This solution of this system is given by

$$\begin{aligned} x_1 &= \frac{1}{6}(4\gamma - 1), \\ x_2 &= \frac{1}{6}(2\gamma + 1), \\ \lambda &= \frac{1}{6}(-8\gamma - 4). \end{aligned}$$

For this solution to be feasible, we require λ to be positive, which is guaranteed if $\gamma < -1/2$. Expressing the solution as a function of γ yields

$$\mathbf{x}^* = \begin{cases} \frac{1}{6}(4\gamma - 1, 2\gamma + 1) & \text{if} \quad \gamma < -\frac{1}{2}, \\ (-\frac{1}{2}, 0) & \text{if} \quad \gamma \geq -\frac{1}{2}. \end{cases}$$

6, B

(b) (i) We begin by defining the Lagrangian

meth seen ⇓

$$\begin{aligned} L(\mathbf{x}, \mathbf{w}, \mu) &= \|\mathbf{x} - \mathbf{y}\|^2 + \lambda \|\mathbf{w}\|_1 + \mu^\top(\mathbf{L}\mathbf{x} - \mathbf{w}), \\ &= \|\mathbf{x} - \mathbf{y}\|^2 + (\mathbf{L}\mu)^\top \mathbf{x} + \lambda\|\mathbf{w}\|_1 - \mu^\top \mathbf{w}. \end{aligned}$$

The Lagrangian is separable with respect to $\mathbf{x}$ and $\mathbf{w}$ and thus we can perform the minimization separately. The minimum of

$$\|\mathbf{x} - \mathbf{y}\|^2 + \left(\mathbf{L}^\top \mu\right)^\top \mathbf{x}$$

over $\mathbf{x}$ is attained when the gradient vanishes,

$$2(\mathbf{x}-\mathbf{y})+\mathbf{L}^\top\mu=0\,,$$

and hence $\mathbf{x}=\mathbf{y}-\frac{1}{2}\mathbf{L}^\top\mu$. Substituting this value back to the $\mathbf{x}$ -part of the Lagrangian, we obtain

$$\min_{\mathbf{x}}\|\mathbf{x}-\mathbf{y}\|^2+\left(\mathbf{L}^\top\mu\right)^\top\mathbf{x}=-\frac{1}{4}\mu^\top\mathbf{L}\mathbf{L}^\top\mu+\mu^\top\mathbf{L}\mathbf{y}$$

In addition,

$$\min_{\mathbf{w}}\lambda\|\mathbf{w}\|_1-\mu^\top\mathbf{w}=\begin{cases}0, & \|\mu\|_\infty\leq\lambda\\ -\infty & \text{else.}\end{cases}$$

To conclude, the dual objective function is given by

$$q(\mu)=\min_{\mathbf{x},\mathbf{w}}L(\mathbf{x},\mathbf{w},\mu)=\begin{cases}-\frac{1}{4}\mu^\top\mathbf{L}\mathbf{L}^\top\mu+\mu^\top\mathbf{L}\mathbf{y}, & \|\mu\|_\infty\leq\lambda\\ -\infty & \text{else.}\end{cases}$$

Therefore, the dual problem is

$$\begin{array}{ll}\max & -\frac{1}{4}\mu^\top\mathbf{L}\mathbf{L}^\top\mu+\mu^\top\mathbf{L}\mathbf{y}\\ \text{s.t.} & \|\mu\|_\infty\leq\lambda\,.\end{array}$$

$\boxed{6,\ \mathrm{C}}$

(ii) Since the primal problem is a convex minimization with an affine constraint, strong duality holds, and primal and dual problems achieve the same value. Therefore, solving the dual problem leads to the solution of the primal via $\mathbf{x}=\mathbf{y}-\frac{1}{2}\mathbf{L}^\top\mu$. In this case, the primal problem is non-smooth because of the ℓ_1 penalization, while the dual problem is the maximization of a quadratic function over a box. In this case, setting a projected gradient ascent/descent is straightforward, as we can use the formula for the orthogonal projection onto a box

$$\mathbf{y}=P_{[-\lambda,\lambda]^n}(\mathbf{x})\,,\quad y_i=\begin{cases}\lambda & x_i\geq\lambda\,,\\ x_i & |x_i|<\lambda\,,\\ -\lambda & x_i\leq-\lambda\,.\end{cases}$$

$\boxed{4,\ \mathrm{B}}$

5. (a) The Hamiltonian is given by

$$H(t,x,u,\lambda) = x_2 + u^2 + \lambda_1(-x_1+u) + \lambda_2 x_1$$

Pointwise minimization is obtained via

$$0 = \frac{\partial H}{\partial u}(t,x,u,\lambda) = 2u + \lambda_1 \quad \Rightarrow \quad u^* = -\frac{1}{2}\lambda_1$$

since H is strictly convex in u. The adjoint equations are given by

$$\begin{aligned}\dot{\lambda}_1(t) &= -\frac{\partial H}{\partial x_1}\left(t, x(t), \mu^*(t,x(t),\lambda(t)), \lambda(t)\right) = \lambda_1(t) - \lambda_2(t)\\ \dot{\lambda}_2(t) &= -\frac{\partial H}{\partial x_2}\left(t, x(t), \mu^*(t,x(t),\lambda(t)), \lambda(t)\right) = -1\end{aligned}$$

with boundary conditions ($\phi(T,x(T)) \equiv 0$)

$$\lambda(T) = \frac{\partial \phi}{\partial x}(T,x(T)) \quad \Longleftrightarrow \quad \lambda_1(T) = \lambda_2(T) = 0$$

Thus, we get $\lambda_2(t) = T - t$ and

$$\dot{\lambda}_1(t) = \lambda_1(t) + t - T, \quad \lambda_1(T) = 0$$

which has the solution

$$\lambda_1(t) = -(t-T) - 1 + e^{t-T}$$

and the optimal control is

$$u^*(t) = -\frac{1}{2}\lambda_1(t) = \frac{1}{2}\left(1 + t - T - e^{t-T}\right).$$

$\boxed{10, \text{M}}$

(b) For this problem, the Hamiltonian is given by

$$H(x,u,\lambda) = \frac{1}{2}u^2 + \frac{1}{2}x^2 + x^4 + \lambda\left(x^3 + u\right).$$

Then, point-wise optimization yields

$$\begin{aligned}\tilde{\mu}(x,\lambda) &\triangleq \arg\min_{u\in\mathbb{R}} H(x,u,\lambda)\\ &= \arg\min_{u\in\mathbb{R}} \left\{\frac{1}{2}u^2 + \frac{1}{2}x^2 + x^4 + \lambda\left(x^3+u\right)\right\} = -\lambda,\end{aligned}$$

and the optimal control is

$$u^*(t) \triangleq \tilde{\mu}\left(x(t), V_x(x(t))\right) = -V_x(x(t))$$

where $V_x(x(t))$ is obtained by solving the infinite time horizon HJB PDE

$$0 = H\left(x, \tilde{\mu}\left(x, V_x\right), V_x\right).$$

In our case, this is equivalent to

$$0 = -\frac{1}{2}V_x^2 + \frac{1}{2}x^2 + x^4 + V_x x^3.$$

We make the ansatz $V(x) = \alpha x^2 + \beta x^4$ which gives $V_x = 2\alpha x + 4\beta x^3$ and we have

$$\begin{aligned}
0 &= -\frac{1}{2}\left(2\alpha x + 4\beta x^3\right)^2 + \frac{1}{2}x^2 + x^4 + \left(2\alpha x + 4\beta x^3\right) x^3 \\
&= -2\alpha^2 x^2 - 8\alpha\beta x^4 - 8\beta^2 x^6 + \frac{1}{2}x^2 + x^4 + 2\alpha x^4 + 4\beta x^6 \\
&= \left(-2\alpha^2 + \frac{1}{2}\right) x^2 + (-8\alpha\beta + 1 + 2\alpha)x^4 + \left(-8\beta^2 + 4\beta\right) x^6
\end{aligned}$$

which is true for all x if $\alpha = \frac{1}{2}$ and $\beta = \frac{1}{2}$ and we have

$$V(x) = \frac{1}{2}\left(x^2 + x^4\right)$$

Finally, the optimal control is given by

$$u^*(t) = -V_x(x(t)) = -x(t) - 2x^3(t)\,.$$

$\boxed{10, \text{M}}$

Review of mark distribution:

Total A marks: 34 of 32 marks

Total B marks: 20 of 20 marks

Total C marks: 10 of 12 marks

Total D marks: 16 of 16 marks

Total marks: 100 of 80 marks

Total Mastery marks: 20 of 20 marks

If your module is taught across multiple year levels, you might have received this form for each level of the module. You are only required to fill this out once for each question.

Please record below, some brief but non-trivial comments for students about how well (or otherwise) the questions were answered. For example, you may wish to comment on common errors and misconceptions, or areas where students have done well. These comments should note any errors in and corrections to the paper. These comments will be made available to students via the MathsCentral Blackboard site and should not contain any information which identifies individual candidates. Any comments which should be kept confidential should be included as confidential comments for the Exam Board and Externals. If you would like to add formulas, please include a separate pdf file with your email.

ExamModuleCode	QuestionN	Comments for Students
Optimisation_MATH60005 MATH97405 MATH70005	1	In general the students handled this question well. Most errors were the incorrect computation of the second derivative of the function in part a to classify the stationary point and, surprisingly, incorrectly identifying that maxima are associated with det(Q) < 0.
Optimisation_MATH60005 MATH97405 MATH70005	2	Most students did very well on parts a and bi. Most students did not perform the correct analysis for part bi, though there were a good number that identified the inequalities that would need to be used.
Optimisation_MATH60005 MATH97405 MATH70005	3	3ai) was correctly answered by most of the students. For 3aii) many students attempted proving convexity using the definition, which is correct but a bit tedious, since using results of max of convex functions or realizing that it is a distance function to a convex set gives a straightfoward answer. For 3b) a large proportion of the class corectly modified some similar proofs we have done in class (e.g. stationarity over the unit simplex), but many students replicated exactly the same argument without noticing that if the test vector is not feasible.
Optimisation_MATH60005 MATH97405 MATH70005	4	4ai) was ok, most of the students used strict convexity as an argument for uniqueness. In 4aii) many answers correctly computed the KKT conditions, but did not separate between gamma>= -1/2. For 4bi) mots of the answers correctly arrived to the dual formulation (although some intermediate steps were unclear, missing in many cases). In 4bii) instead, not many answers correctly identified the non-differentiability of the primal problem and proposed a concrete solution strategy for the dual, incluing complementary slackness conditions to recover the primal solution.
Optimisation_MATH60005 MATH97405 MATH70005	5	5a) answers were correct in most of the cases, although there were many integration errors when solving the adjoin equation and recovering a solution for the optimal control signal. In 5b), some answers correctly cast the right HJB equation, used the ansatz provided, and arrived to the right solution by matching coefficients. Many incomplete answers. Some attempts didn't use the ansatz, and instead solved a quadratic equation for the gradient of V, which is ok provided solutions that do not correspond to a value function were discarded.