3M1 Mathematical Methods, 2026

1 Optimisation

1. Consider a thin metal rod extending along the x-axis. The temperature distribution $T(x)$ is modeled as the superposition of two heat sources located at $x = 1$ and $x = -1$, where the source at $x = -1$ is twice as powerful as the source at $x = 1$. The parameter $\alpha \in \mathbb{R}$ represents the thermal concentration.

$$T(x) = \exp(-\alpha(x-1)^2) + 2\exp(-\alpha(x+1)^2)$$

(a) i. Derive the necessary condition for a point $x^* \in \mathbb{R}$ to be a critical point of the temperature distribution for any value of α. [10%]

(a) Necessary Condition:
A critical point x^* must satisfy $T'(x^*) = 0$. Using the chain rule:

$$T'(x) = -2\alpha(x-1)e^{-\alpha(x-1)^2} - 4\alpha(x+1)e^{-\alpha(x+1)^2}$$

[10%]

ii. For the case $\alpha = 1$, show that the critical points are defined by the non-linear equation:

$$\frac{1-x}{2(1+x)} = e^{-4x}$$

[10%]

(b) Non-linear Equation for $\alpha = 1$:
Setting $T'(x) = 0$ with $\alpha = 1$:

$$-2(x-1)e^{-(x-1)^2} - 4(x+1)e^{-(x+1)^2} = 0 \implies (1-x)e^{-(x-1)^2} = 2(x+1)e^{-(x+1)^2}$$

Rearranging terms:

$$\frac{1-x}{2(1+x)} = \frac{e^{-(x+1)^2}}{e^{-(x-1)^2}} = e^{-(x+1)^2+(x-1)^2} = e^{-4x}$$

[10%]

iii. For the necessary condition use a second-order Taylor expansion of $T(x)$ to derive the Newton iteration scheme to find x^*. [15%]

(c) Newton Iteration Scheme:
Expanding $T(x)$ as a second-order Taylor series around x_k:

$$T(x_k + \Delta x) \approx T(x_k) + T'(x_k)\Delta x + \frac{1}{2}T''(x_k)\Delta x^2$$

Stationary point of the quadratic approximation occurs at $\Delta x = -T'(x_k)/T''(x_k)$.
Update rule:

$$x_{k+1} = x_k - \frac{T'(x_k)}{T''(x_k)}$$

Using the product rule for $T''(x)$:

$$T''(x) = -2\alpha \left[e^{-\alpha(x-1)^2}(1 - 2\alpha(x-1)^2) + 2e^{-\alpha(x+1)^2}(1 - 2\alpha(x+1)^2) \right]$$

[15%]

(b) i. Derive and justify the sufficient condition for a point $x^* \in \mathbb{R}$ to be a **strong local maximum** (a local hot spot). [10%]

(a) Sufficient Condition:

A strong local maximum requires $T'(x^*) = 0$ and $T''(x^*) < 0$. Negative curvature ensures the stationary point is a peak (concave down). [10%]

ii. For $\alpha = 1$, explain, without solving the non-linear equation, which of the non-zero critical points corresponds to the **global maximum** temperature. [10%]

(b) Global Maximum Identification:

The source at $x = -1$ has double the amplitude (weight 2) of the source at $x = 1$. Since Gaussian interference is additive and positive, the peak near the stronger source ($x \approx -1$) is the global maximum. [10%]

iii. For $\alpha = 1$, using the derived Newton iteration scheme, perform one manual iteration starting from the stronger source center to demonstrate whether the global maximum lies to the left or right of the source center, and provide a physical justification for this shift. [15%]

(c) Manual Iteration and Shift Analysis:

Start at $x_0 = -1$. Let $R(x) = e^{-4x}$.

$$N = (x_0 - 1) + 2(x_0 + 1)R(x_0) = (-2) + 2(0)e^4 = -2$$

$$D = (1-2(x_0-1)^2)+2R(x_0)(1-2(x_0+1)^2) = (1-8)+2e^4(1-0) = -7+109.20 = 102.20$$

$$\Delta x = N/D = -2/102.20 \approx -0.0196 \implies x_1 = -1 - (-0.0196) = \mathbf{-0.9804}$$

Interpretation: $x_1 > x_0$, so the peak shifted **right**. This is due to the positive gradient from the source at $x = 1$ adding to the temperature at $x = -1$. [15%]

(c) i. Determine the range of values for α for which the temperature distribution $T(x)$ is **convex** for all $x \in \mathbb{R}$. [20%]

(a) Range of α:

$T(x)$ is convex if $T''(x) \geq 0$.

- For $\alpha > 0$, $T''(1) \approx -2\alpha < 0$ (for large α), so not convex.
- For $\alpha \leq 0$, let $\beta = -\alpha \geq 0$. $T(x) = e^{\beta(x-1)^2} + 2e^{\beta(x+1)^2}$. Both terms are convex functions ($\beta \geq 0$). Sum of convex functions is convex.

Range: $\alpha \leq 0$. [20%]

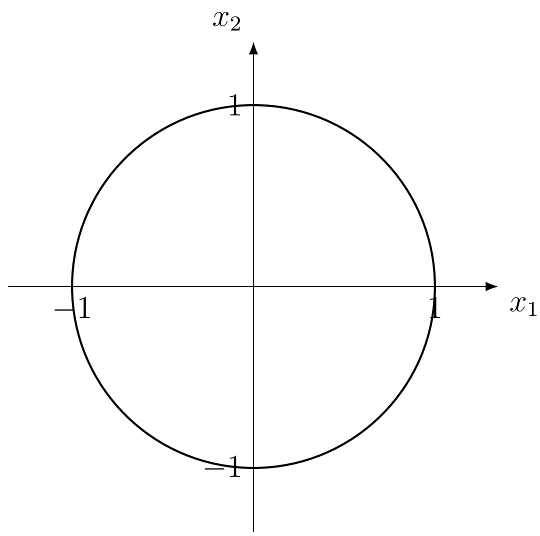
ii. State the implication of this convexity on the nature of any critical point x^*. [10%]

(b) Implication:

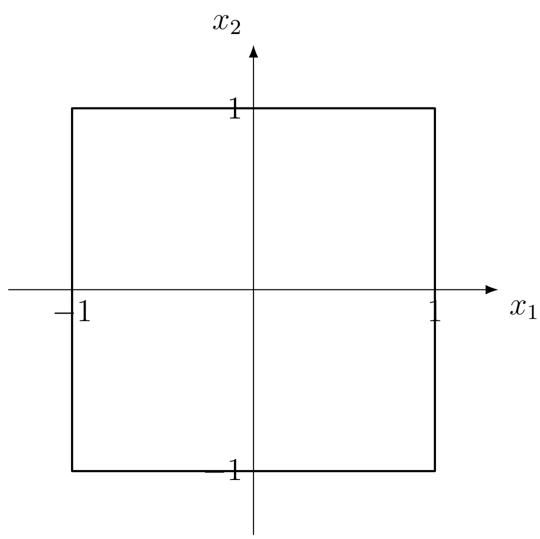
On a convex domain, any critical point x^* is a **global minimum**. [10%]

2 Linear Algebra

1. (a) i. • ℓ_2-norm:



• ℓ_∞-norm:



[10%]

ii. Let $x \in \mathbb{C}^n$. Let k be the index of a component of x with maximum magnitude. We have

$$\|x\|_2^2 = \sum_{i=1}^{n} |x_i|^2 \geq |x_k|^2 = \|x\|_\infty^2.$$

We also have

$$\|x\|_2^2 = \sum_{i=1}^{n} |x_i|^2 \leq \sum_{i=1}^{n} |x_k|^2 = n\|x\|_\infty^2.$$

Hence,

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\,\|x\|_\infty. \tag{1}$$

[10%]

iii. Now,

$$\|Ax\|_2 \leq \sqrt{n}\,\|Ax\|_\infty \leq \sqrt{n}\,\|A\|_\infty\,\|x\|_\infty \leq \sqrt{n}\,\|A\|_\infty\,\|x\|_2.$$

Divide by $\|x\|_2$ and take the maximum over $x \neq 0$ to obtain

$$\|A\|_2 \leq \sqrt{n}\,\|A\|_\infty.$$

Again by (1),

$$\|Ax\|_\infty \leq \|Ax\|_2 \leq \|A\|_2\,\|x\|_2 \leq \|A\|_2\,\sqrt{n}\,\|x\|_\infty.$$

Divide by $\|x\|_\infty$ and take the maximum over $x \neq 0$ to obtain

$$\|A\|_\infty \leq \sqrt{n}\,\|A\|_2,$$

Hence, we have

$$\boxed{\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{n}\|A\|_\infty}$$

with $c = \frac{1}{\sqrt{n}}$ and $C = \sqrt{n}$.

[20%]

(b) i. Geometrically, the map $x \mapsto Ax$ is the composition of three simple transformations:

$$x \xmapsto{V^H} V^H x \xmapsto{\Sigma} \Sigma V^H x \xmapsto{U} U\Sigma V^H x.$$

- Since V is unitary, V^H is also unitary, so it preserves lengths and angles. In $\mathbb{R}^n$, it rotates/reflects the unit sphere onto itself, such that the right singular vectors (columns of V) align with the coordinate axes.
- Σ is diagonal with nonnegative entries, so it stretches by a factor σ_i in the ith coordinate direction and collapses directions corresponding to $\sigma_i = 0$. In particular, the unit sphere is mapped to an axis-aligned ellipsoid with semi-axis lengths $\sigma_1, \ldots, \sigma_n$.
- U is unitary, so it preserves lengths and angles and rotates/reflects the ellipsoid such that its principal axes align with the left singular vectors (columns of U).

[20%]

ii. The polar decomposition of A splits it into two "simple" parts: a positive-semidefinite "stretch" P that scales along orthogonal directions and a unitary "rotation" Q.

[5%]

iii. Let $A \in \mathbb{C}^{n\times n}$ have SVD $A = U\Sigma V^H$ and let $Q = UV^H$ as in part (b). Let $W \in \mathbb{C}^{n\times n}$ be any unitary matrix. Since the Frobenius norm is unitarily invariant, we have

$$\begin{aligned}\|A - Q\|_F &= \|U\Sigma V^H - UV^H\|_F \\ &= \|U^H(U\Sigma V^H - UV^H)V\|_F \\ &= \|\Sigma - I\|_F.\end{aligned}$$

Similarly,

$$\begin{aligned}\|A - W\|_F &= \|U\Sigma V^H - W\|_F \\ &= \|U^H(U\Sigma V^H - W)V\|_F \\ &= \|\Sigma - U^H W V\|_F \\ &= \|\Sigma - Z\|_F,\end{aligned}$$

where $Z := U^H W V$ is unitary (as a product of unitary matrices).
Thus we need to show that for every unitary Z,

$$\|\Sigma - I\|_F \leq \|\Sigma - Z\|_F.$$

Using

$$\|B - C\|_F^2 = \|B\|_F^2 + \|C\|_F^2 - 2\,\Re\Big(\operatorname{tr}\Big(C^H B\Big)\Big).$$

given in the question, we have

$$\|\Sigma - Z\|_F^2 = \|\Sigma\|_F^2 + n - 2\,\Re\Big(\operatorname{tr}\Big(Z^H \Sigma\Big)\Big)$$

and

$$\|\Sigma - I\|_F^2 = \|\Sigma\|_F^2 + n - 2\,\Re\big(\operatorname{tr}(\Sigma)\big).$$

So we need to show that

$$\Re\big(\operatorname{tr}(\Sigma Z)\big) \leq \operatorname{tr}(\Sigma).$$

We have $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_i \geq 0$. Then

$$\Re\big(\operatorname{tr}(\Sigma Z)\big) = \sum_{i=1}^{n} \sigma_i\,\Re(z_{ii}).$$

For a unitary matrix, each diagonal entry satisfies $|z_{ii}| \leq 1$ (since each column has Euclidean norm 1), hence $\Re(z_{ii}) \leq |z_{ii}| \leq 1$. Therefore

$$\Re\big(\operatorname{tr}(\Sigma Z)\big) = \sum_{i=1}^{n} \sigma_i\,\Re(z_{ii}) \leq \sum_{i=1}^{n} \sigma_i = \operatorname{tr}(\Sigma).$$

Hence

$$\boxed{\|A - Q\|_F \leq \|A - W\|_F \quad \text{for all unitary } W.}$$

This result says that, among all unitary matrices, $Q = UV^H$ is a closest unitary matrix to A in Frobenius norm. [35%]

3 Stochastic Process

1. *Self-Normalising Importance Sampling*

(a)(i) This is a standard Monte-Carlo approximation so

$$\hat{\mu}_N = \frac{1}{N}\sum_{i=1}^{N} h(\boldsymbol{x}^{(i)})$$

[10%]

(a)(ii) Initially compute the expected value of the integral estimate from a single sample drawn from $p(\boldsymbol{x})$. To show the expected value of the estimate is unbiased

$$\mathbb{E}\left\{\hat{\mu}_N\right\} = \mathbb{E}\left\{\frac{1}{N}\sum_{i=1}^{N} h(\boldsymbol{x}^{(i)})\right\} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left\{h(\boldsymbol{x})\right\} = \mu$$

for all values of N. Now computing the variance

$$\mathrm{var}(\hat{\mu}_1) = \mathbb{E}\left\{(h(\boldsymbol{x}) - \mu)^2\right\} = \sigma^2$$

from the question. The variance from N independent estimators is then given by

$$\mathrm{var}(\hat{\mu}_N) = \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{var}(\hat{\mu}_1) = \frac{\sigma^2}{N}$$

The variance does not depend on the dimensionality of $\boldsymbol{x}$, just the number of samples N and the variance σ^2. [30%]

(b)(i) Consider the expected value of an element the denominator

$$\frac{1}{N}\sum_{i=1}^{N}\frac{p^\star(\tilde{\boldsymbol{x}}^{(i)})}{q^\star(\tilde{\boldsymbol{x}}^{(i)})} \approx \int \frac{p^\star(\boldsymbol{x})}{q^\star(\boldsymbol{x})} q(\boldsymbol{x}) d\boldsymbol{x} = \int \frac{p^\star(\boldsymbol{x})}{Z_{\mathtt{q}}} d\boldsymbol{x} = \frac{Z_{\mathtt{p}}}{Z_{\mathtt{q}}}$$

Now consider the numerator

$$\frac{1}{N}\sum_{i=1}^{N} h(\tilde{\boldsymbol{x}}^{(i)})\frac{p^\star(\tilde{\boldsymbol{x}}^{(i)})}{q^\star(\tilde{\boldsymbol{x}}^{(i)})} \approx \int h(\boldsymbol{x})\frac{p^\star(\boldsymbol{x})}{q^\star(\boldsymbol{x})} q(\boldsymbol{x}) d\boldsymbol{x} = \int h(\boldsymbol{x})\frac{p^\star(\boldsymbol{x})}{Z_{\mathtt{q}}} dx = V\frac{Z_{\mathtt{p}}}{Z_{\mathtt{q}}}$$

So dividing through yields the required integral

$$\tilde{\mu}_N \approx \frac{\sum_{i=1}^{N} h(\tilde{\boldsymbol{x}}^{(i)})\frac{p^\star(\tilde{\boldsymbol{x}}^{(i)})}{q^\star(\tilde{\boldsymbol{x}}^{(i)})}}{\sum_{j=1}^{N}\frac{p^\star(\tilde{\boldsymbol{x}}^{(j)})}{q^\star(\tilde{\boldsymbol{x}}^{(j)})}}, \quad w^{(i)} = \frac{p^\star(\tilde{\boldsymbol{x}}^{(i)})}{q^\star(\tilde{\boldsymbol{x}}^{(i)})}$$

The underlying assumption here is that $q(x)$ is non-zero for all non-zero values of $p(x)$, otherwise in the limit the correct value of V will not be obtained.

Note though it is not possible to compute Z_q it is possible to draw samples from $q(x)$ as the normalisation term is requiured to draw samples. [35%]

(b)(ii) For this expression to converge to μ it is necessary for $q(\boldsymbol{x})$ to be non-zero wherever $p(\boldsymbol{x})$ is non-zero. Otherwise there will be a region of the space that cannot contribute to the calculation of μ. [10%]

(b)(iii) The estaimate is consistent, not unbiased. The issue is that the estimate can be written as a ratio of random variables and the following is true

$$\mathbb{E}\left\{\frac{\mu}{w}\right\} \neq \frac{\mathbb{E}\left\{\mu\right\}}{\mathbb{E}\left\{w\right\}}$$

As $N \to \infty$ the variances in the terms (under mild conditions) will go towards zero, so the estimate is consistent. [15%]