MAMA/326, NST3AS/326, MAAS/326

MAT3

MATHEMATICAL TRIPOS Part III

Monday 9 June 2025 1:30 pm to 4:30 pm

PAPER 326

INVERSE PROBLEMS

Before you begin please read these instructions carefully

Candidates have THREE HOURS to complete the written examination.

Attempt **ALL** questions.
There are **THREE** questions in total.
The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury tag
Script paper
Rough paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1 Variational regularisation

Suppose that $\mathcal{X}, \mathcal{Y}$ are Banach spaces, that $\mathcal{X}$ is reflexive, and that we have a bounded linear forward operator $A : \mathcal{X} \to \mathcal{Y}$ mapping images u to measurements f. In the lectures, we studied optimisation problems of the following form to obtain approximate solutions to the corresponding inverse problem:

$$\hat{u}_\alpha \in \operatorname*{argmin}_{u \in \mathcal{X}} \frac{1}{2}\|Au - f\|_{\mathcal{Y}}^2 + \alpha \mathcal{J}(u), \tag{1}$$

for measurements f, and with $\mathcal{J} : \mathcal{X} \to \mathbf{R} \cup \{\infty\}$ a proper, convex, lower semicontinuous regularisation functional that is bounded from below, and such that the objective function in (1) is coercive.

(a) Explain why existence of solutions to (1) is guaranteed under the stated conditions.

(b) Given $f \in \mathcal{Y}$ and $u^\dagger \in \mathcal{X}$, state what it means for $u^\dagger$ to be a $\mathcal{J}$-minimising solution to the inverse problem $Au = f$. Given such a $u^\dagger$, state what it means for $u^\dagger$ to satisfy the source condition.

(c) Let us now consider a modification of the standard variational problem (1), given as follows:

$$\hat{u}_\alpha \in \operatorname*{argmin}_{u \in \mathcal{X}} \|Au - f\|_{\mathcal{Y}} + \alpha \mathcal{J}(u). \tag{2}$$

Suppose further that there is a $\mathcal{J}$-minimising solution $u^\dagger$ satisfying the source condition.
Calculate an $0 < \alpha_0 \leqslant \infty$, depending on the element in the source condition, such that any solution $\hat{u}_\alpha$ to (2) is a $\mathcal{J}$-minimising solution for $0 < \alpha < \alpha_0$. If additionally $\mathcal{J}$ is strictly convex, conclude that the solution $\hat{u}_\alpha$ to (2) coincides exactly with $u^\dagger$ for $0 < \alpha < \alpha_0$.
Hint: use that $\hat{u}_\alpha$ satisfies (2) and compare to an appropriate other element of $\mathcal{X}$.

(d) Of course, in practical settings, we do not have access to the clean measurements f and instead have to make do with noisy measurements f_δ, which we assume satisfy $\|f - f_\delta\|_{\mathcal{Y}} \leqslant \delta$. In this case, (2) is modified to use f_δ:

$$\hat{u}_{\alpha,\delta} \in \operatorname*{argmin}_{u \in \mathcal{X}} \|Au - f_\delta\|_{\mathcal{Y}} + \alpha \mathcal{J}(u). \tag{3}$$

We will continue to assume that there is a $\mathcal{J}$-minimising solution $u^\dagger$ satisfying the source condition.
Recall the definition of the Bregman divergence of a convex functional and show the following bound on the Bregman divergence of $\mathcal{J}$ (with $p^\dagger$ the specific subgradient from the definition of the source condition)

$$D_{\mathcal{J}}^{p^\dagger}(\hat{u}_{\alpha,\delta}, u^\dagger) \leqslant C_\alpha \delta,$$

for any α satisfying the same constraints $0 < \alpha < \alpha_0$ as in the previous step. Here $C_\alpha \geqslant 0$ depends on α (and on the element in the source condition) but not on δ.
Comment: with this formulation, the inequality above shows that we do not need to let α converge to 0 as $\delta \to 0$ to get convergence in the Bregman divergence, unlike in the results in the lectures!

[QUESTION CONTINUES ON THE NEXT PAGE]

(e) Under the same assumptions, we can consider a different but related approach to variational regularisation through constrained optimisation. We will let $c \geqslant 1$ be arbitrary and solve the following problem to get an approximate inverse:

$$\hat{u}_\delta \in \operatorname*{argmin}_{u \in E_\delta} \mathcal{J}(u), \qquad \text{where} \qquad E_\delta := \{u \in \mathcal{X} \,|\, \|Au - f_\delta\|_\mathcal{Y} \leqslant c\delta\}. \tag{4}$$

Explain why solutions to (4) exist, using a similar reasoning as in (a). Under the source condition, as above, prove the following bound on the Bregman divergence:

$$D_\mathcal{J}^{p^\dagger}(\hat{u}_\delta, u^\dagger) \leqslant C\delta$$

for some $C \geqslant 0$ only depending on c and the element in the source condition.

2 Bayesian inverse problems

Suppose that $\mathcal{X} = L^2[0,1]$. In what follows, when we speak of measures on $\mathcal{X}$, we will always mean that they are defined on the measurable space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. We will further assume the existence of an underlying probability space $(\Omega, \mathcal{F}, \mathbf{P})$ on which we have a sequence $\{\xi_i\}_{i=0}^{\infty}$ of independent standard Gaussian random variables. We are given a problem in which an initial condition $u \in \mathcal{X}$ is diffused according to the heat equation

$$\begin{cases} \frac{\partial v}{\partial t}(x,t) = \frac{\partial^2 v}{\partial x^2}(x,t) & \text{on} \quad [0,1] \times \mathbf{R}_{\geqslant 0} \\ \frac{\partial v}{\partial x}(x,t) = 0, & \text{on} \quad \{0,1\} \times \mathbf{R}_{>0}, \\ v(x,0) = u(x) & \text{on} \quad [0,1], \end{cases}$$

and we have $N \geqslant 1$ uniformly spaced sensors at $x_i := (i+1)/(N+2) \in (0,1)$ for $i = 1, \ldots, N$. With these sensors, we can take noisy measurements of point evaluations $\{v_i\}_{i=1}^N := \{v(x_i, 1)\}_{i=1}^N$, and let us write $\mathcal{G} : \mathcal{X} \to \mathbf{R}^N$ for the map $u \mapsto \{v(x_i,1)\}_{i=1}^N$. You may use without proof that for any $t > 0$, $u \mapsto v(\cdot, t)$ is a continuous map from $L^2[0,1] \to C[0,1]$ (in particular, point evaluations of v are well defined for $t > 0$). We will consider the inverse problem of recovering u from $\{\tilde{v}_i\}_{i=1}^N$, a noisy version of $\{v_i\}_{i=1}^N$ and frame it as a Bayesian inverse problem.

(a) Consider the integral operator K defined by

$$Kf(x) = \int_0^1 \exp(-|x-y|) f(y) \, \mathrm{d}y.$$

Show that K is a self-adjoint, bounded linear operator $\mathcal{X} \to \mathcal{X}$.

(b) Show that the eigenvalues λ of K satisfy the equation

$$\frac{2\omega}{\omega^2 - 1} = \tan(\omega), \tag{1}$$

where $-\omega^2 = (\lambda - 2)/\lambda$.
Hint: The eigendecomposition can be found by solving an integral equation. Derive an equivalent ODE with boundary conditions and study its solutions.

(c) Without explicitly solving the transcendental equation (1), show that the corresponding eigenvalues satisfy $\lambda_n > 0$ and $\lambda_n = \mathcal{O}(1/n^2)$ as $n \to \infty$.

(d) In fact, K is compact (you do not need to prove this).
Using the results above, conclude that K can be used as the covariance operator of a Gaussian random field: how would you construct a random variable taking values in $\mathcal{X}$, the law of which is the measure $\mathrm{N}(m, K)$? You may use results from the lectures without proof as long as they are clearly stated.

(e) We will model the initial condition as a random variable $U : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to (\mathbf{R}, \mathcal{B}(\mathbf{R}))$, taking the prior measure to be $\mu_0 = \mathbf{P}(U \in \cdot) := \mathrm{N}(m, K)$ for some $m \in L^2[0,1]$, and we will model the noisy measurements as a random variable $V : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to (\mathbf{R}^N, \mathcal{B}(\mathbf{R}^N))$ with conditional law $\mathbf{P}(V \in \cdot | U = u) = \mathrm{N}(\mathcal{G}(u), \sigma^2 \, \mathrm{id}_{\mathbf{R}^N})$ for some $\sigma > 0$ representing the noise level.
What is the corresponding likelihood? State the Bayesian inverse problem of recovering U from V.

[QUESTION CONTINUES ON THE NEXT PAGE]

(f) Recall the definition of the total variation distance, $d_{\mathrm{TV}}(\mu, \nu)$, between two probability measures $\mu, \nu \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. What does it mean for the Bayesian inverse problem to be $(\mathrm{Prob}(\mathcal{X}, \mathcal{B}(\mathcal{X})), d_{\mathrm{TV}})$-well-posed? Show that the Bayesian inverse problem defined in (c) is well-posed in this sense. You may use results from the lectures without proof as long as they are clearly stated.

3 Classical regularisation

We will consider a linear inverse problem with forward operator $A : \mathcal{X} \to \mathcal{Y}$ between infinite-dimensional Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$:

$$\text{Estimate } u \text{ from } f \text{ where } Au = f. \tag{1}$$

We will further assume that A is a *compact operator* with infinite-dimensional range $\mathrm{im}(A)$.

(a) What does it mean for the inverse problem, (1), to be well-posed in the sense of Hadamard? Given the assumptions on A, why is the inverse problem *always* ill-posed?

(b) Given a bounded linear operator $K : \mathcal{X} \to \mathcal{X}$, prove that

$$(\mathrm{id}_{\mathcal{X}} - K)^{-1} = \sum_{n=0}^{\infty} K^n, \tag{2}$$

if $\|K\|_{\mathcal{L}(\mathcal{X},\mathcal{X})} < 1$, where the series converges with respect to the operator norm. As a consequence, show that, if $\tau \neq 0$ and $\|\mathrm{id}_{\mathcal{X}} - \tau K\|_{\mathcal{L}(\mathcal{X},\mathcal{X})} < 1$, then

$$K^{-1} = \tau \sum_{n=0}^{\infty} (\mathrm{id}_{\mathcal{X}} - \tau K)^n. \tag{3}$$

(c) Recall the normal equation associated with (1). Under what conditions does the normal equation have solutions, and under what conditions are these solutions unique? How do the solutions of the normal equation relate to the Moore–Penrose inverse?

(d) It is tempting to try to apply (3) with $K = A^*A$ to solve the normal equation. Why is this not possible, even if $\ker(A) = \{0\}$?
Hint: apply the spectral theorem to A^*A to show that the conditions we assumed to derive (3) *are not satisfied.*

(e) Despite the negative result in (d), the series in (3) can be used to compute solutions to the normal equation when they exist:
Recall the domain, $\mathrm{dom}(A^\dagger)$, of the Moore–Penrose inverse and show that if $f \in \mathrm{dom}(A^\dagger)$ and τ is chosen so that $0 < \tau < 1/\|A\|^2_{\mathcal{L}(\mathcal{X},\mathcal{Y})}$, then

$$Q_N f = \tau \sum_{n=0}^{N} (\mathrm{id}_{\mathcal{X}} - \tau A^*A)^n A^* f \to A^\dagger f, \tag{4}$$

as $N \to \infty$.
Hint: it may be useful to recall the Picard criterion.

(f) Show that the series in (4) can be used to define a convergent regularisation of the inverse problem (1), i.e., show that there is an a-priori parameter choice rule $\alpha : \mathbf{R}_{>0} \to \mathbf{R}_{>0}$ and a regularisation $\{R_\alpha\}_{\alpha>0}$ defined using this series such that the pair $(\{R_\alpha\}_{\alpha>0}, \alpha)$ is a convergent regularisation. You may use a result proven in the lectures, provided you state it clearly.

END OF PAPER