Inverse Problems — Summary

Lucas Riedstra

December 8, 2020

Contents

1 Introduction

TODO: write this section (plan is to write this at the end of the course with new knowledge :-)).

A *direct problem* is a problem where given an object or *cause*, we must determine the data or *effect*. In an *inverse problem*, we observe data and wish to recover the object.

A problem is called *well-posed* if a unique solution exists that depends continuously on the data. Most inverse problems are, unfortunately, ill-posed.

2 Generalised Solutions

Recap 2.1. 1. An linear operator $A\colon \mathcal{X} \to \mathcal{Y}$ is called *bounded* if

$$\|A\|_{\mathcal{B}(\mathcal{X},\mathcal{Y})} \coloneqq \sup_{u\neq 0} \frac{\|Au\|_{\mathcal{Y}}}{\|u\|_{\mathcal{X}}} = \sup_{\|u\|_{\mathcal{X}}\leq 1} < \infty.$$

It is known that a linear operator between normed spaces is continuous if and only if it is bounded. The set of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ is denotes $\mathcal{B}(\mathcal{X},\mathcal{Y})$.

2. We let $\mathcal{D}(A)$, $\mathcal{N}(A)$ and $\mathcal{R}(A)$ denote the domain, null space, and range of A respectively.

3. We will assume $\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces, so there is an inner product $\langle\cdot,\cdot\rangle$ and any bounded operator $A \in \mathcal{B}(\mathcal{X},\mathcal{Y})$ has a unique adjoint $A^* \in \mathcal{B}(\mathcal{Y},\mathcal{X})$ which satisfies

$$\langle Au, v\rangle_{\mathcal{Y}} = \langle u, A^*v\rangle_{\mathcal{X}} \quad \text{for all } u \in \mathcal{X},\, v \in \mathcal{Y}.$$

4. For any $\mathcal{X}' \subseteq \mathcal{X}$ we define the *orthogonal complement* of $\mathcal{X}'$ as

$$(\mathcal{X}')^\perp \coloneqq \{u \in \mathcal{X} \mid \langle u, v\rangle_{\mathcal{X}} = 0 \ \forall v \in \mathcal{X}'\}.$$

It is known that $(\mathcal{X}')^\perp$ is a closed subspace of $\mathcal{X}$ and that $\mathcal{X}' \subseteq ((\mathcal{X}')^\perp)^\perp$, where equality holds if and only if $\mathcal{X}'$ is a closed subspace of $\mathcal{X}$. For a non-closed subspace $\mathcal{X}'$ we have $((\mathcal{X}')^\perp)^\perp = \overline{\mathcal{X}'}$.

5. If $\mathcal{X}'$ is a closed subspace of $\mathcal{X}$, then for any $u \in \mathcal{X}$ there exist unique $x_u \in \mathcal{X}'$, $x_u^\perp \in (\mathcal{X}')^\perp$ such that $u = x_u + x_u^\perp$. The map $u \mapsto x_u$ is denoted $P_{\mathcal{X}'}$ and is called the *orthogonal projection* on $\mathcal{X}'$. Properties are:

   (a) $P_{\mathcal{X}'}$ is bounded and self-adjoint with norm 1;

   (b) $P_{\mathcal{X}'} + P_{(\mathcal{X}')^\perp} = I$;

   (c) $P_{\mathcal{X}'}u$ minimises the distance from u to $\mathcal{X}'$;

   (d) $x = P_{\mathcal{X}'}u$ if and only if $x \in \mathcal{X}'$ and $u - x \in (\mathcal{X}')^\perp$.

6. For any $A \in \mathcal{B}(\mathcal{X},\mathcal{Y})$ we have

$$\mathcal{R}(A)^\perp = \mathcal{N}(A^*) \quad \text{and} \quad \mathcal{N}(A)^\perp = \overline{\mathcal{R}(A^*)}.$$

Lemma 2.2. *For any $A \in \mathcal{B}(\mathcal{X},\mathcal{Y})$ we have $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)}$.*

Proof. It is trivial that $\overline{\mathcal{R}(A^*A)} \subseteq \overline{\mathcal{R}(A^*)}$.

Now, suppose $u \in \overline{\mathcal{R}(A^*)}$ and let $\varepsilon > 0$. Then there exists $v \in \mathcal{X}$ such that $\|A^*v - u\| < \varepsilon/2$. Writing $v = e + f$ with $e \in \mathcal{N}(A^*)$, $f \in \mathcal{N}(A^*)^\perp = \overline{\mathcal{R}(A)}$, we see that $\|A^*f - u\| < \varepsilon/2$.

Since $f \in \overline{\mathcal{R}(A)}$, there exists $x \in \mathcal{X}$ such that $\|Ax - f\| < \varepsilon/(2\|A\|)$. We now compute

$$\|A^*Ax - u\| \leq \|A^*Ax - A^*f\| + \|A^*f - u\| < \|A^*\|\frac{\varepsilon}{2\|A\|} + \frac{\varepsilon}{2} = \varepsilon,$$

and conclude that $u \in \overline{\mathcal{R}(A^*A)}$. This shows that $\overline{\mathcal{R}(A)} \subseteq \overline{\mathcal{R}(A^*A)}$. □

2.1 Generalised inverses

We consider the equation

$$Au = f, \tag{1}$$

where $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and f are known, and we wish to find u.

Definition 2.3. An element $u \in \mathcal{X}$ is called a *least-squares solution* of eq. (1) if u is a minimiser of the function $v \mapsto \|Av - f\|_{\mathcal{Y}}$. It is called a *minimal-norm solution* of eq. (1) if it has minimal norm among all least-squares solutions.

Note that a least-squares solution may not exist. If a least-squares solution u exists, then the affine subspace of all least-squares solutions is given by $u + \mathcal{N}(A)$. By writing $u = u^\dagger + v$ for $u^\dagger \in \mathcal{N}(A)^\perp$, $v \in \mathcal{N}(A)$, we find that the space of least-squares solutions is given by $u^\dagger + \mathcal{N}(A)$, and it is now clear that $u^\dagger$ is the unique minimum-norm solution.

Theorem 2.4. *Let $f \in \mathcal{Y}$ and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then the following are equivalent:*

1. *$u \in \mathcal{X}$ satisfies $Au = P_{\overline{\mathcal{R}(A)}} f$;*
2. *u is a least-squares solution of eq. (1):*
3. *u solves the normal equation*

$$A^* f = A^* A u. \tag{2}$$

Proof. "(1) $\implies$ (2)": We have

$$\|Au - f\|_{\mathcal{Y}} = \left\| P_{\overline{\mathcal{R}(A)}} f - f \right\| = \inf_{g \in \overline{\mathcal{R}(A)}} \|g - f\| \leq \inf_{g \in \mathcal{R}(A)} \|g - f\| = \inf_{u \in \mathcal{X}} \|Au - f\|.$$

"(2) $\implies$ (3)": Let $u \in \mathcal{X}$ be a least-squares solution and $v \in \mathcal{X}$ arbitrary. Define the quadratic polynomial

$$\begin{aligned} F \colon \mathbb{R} \to \mathbb{R} \colon \lambda \mapsto & \|A(u + \lambda v) - f\|^2 \\ &= \langle Au + \lambda Av - f, Au + \lambda Av - f \rangle \\ &= \lambda^2 \|Av\|^2 - 2\lambda \langle Av, f - Au \rangle + \|f - Au\|^2. \end{aligned}$$

As u is a least-squares solution, we know that F attains a minimum in $\lambda = 0$ and therefore that

$$0 = F'(0) = 2\langle Av, f - Au \rangle = 2\langle v, A^*(f - Au) \rangle.$$

Since v is arbitrary, we must have $A^*(f - Au) = 0$, so u satisfies *eq.* (2).

"(3) $\implies$ (1)": From the normal equation we know that $A^*(f - Au) = 0$. For any $x \in \mathcal{X}$, we have

$$\langle Ax, f - Au \rangle = \langle x, A^*(f - Au) \rangle = \langle x, 0 \rangle = 0,$$

so $f - Au \in \mathcal{R}(A)^\perp$.

So we have $Au \in \overline{\mathcal{R}(A)}$ and $f - Au \in \mathcal{R}(A)^\perp = \overline{\mathcal{R}(A)}^\perp$, from which it follows that $Au = P_{\overline{\mathcal{R}(A)}} f$. □

The following lemma gives a precise condition for when a least-squares solution exists:

Lemma 2.5. *Equation (1) has a least-squares solution if and only if $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$.*

Proof. " $\implies$ " Suppose u is a least-squares solution. Then $f - Au \in \mathcal{R}(A)^\perp$, so $f = Au + (f - Au) \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$.

" $\impliedby$ " Suppose $f = Au + g$ for some $u \in \mathcal{X}$, $g \in \mathcal{R}(A)^\perp = \overline{\mathcal{R}(A)}^\perp$. Then by the previous theorem, $Au = P_{\overline{\mathcal{R}(A)}} f$, so u is a least-squares solution. □

Corollary 2.6. *If $\mathcal{R}(A)$ is closed, then eq. (1) always has a least-squares solution.*

In particular, this holds if $\mathcal{R}(A)$ is finite-dimensional. Therefore, if either $\mathcal{X}$ or $\mathcal{Y}$ is finite-dimensional, eq. (1) has a least-squares solution for any A.

We have already seen that if a least-squares solution u exists, then the affine subspace of all least-sqares solutions is $u + \mathcal{N}(A)$, and the unique minimum-norm solution is the projection of 0 onto this affine subspace, which is the unique element of $u + \mathcal{N}(A)$ that lies in $\mathcal{N}(A)^\perp$.

Definition 2.7. Let $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, and define

$$\tilde{A} \coloneqq A\restriction_{\mathcal{N}(A)^\perp} \colon \mathcal{N}(A)^\perp \to \mathcal{R}(A).$$

Clearly $\tilde{A}$ is bijective and we define the *Moore-Penrose inverse*

$$A^\dagger \colon \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp \to \mathcal{N}(A)^\perp \colon f \mapsto \tilde{A}^{-1} P_{\overline{\mathcal{R}(A)}} f.$$

Remark. Note that $\overline{\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp} = \overline{\mathcal{R}(A)} \oplus \mathcal{R}(A)^\perp = \overline{\mathcal{R}(A)} \oplus \overline{\mathcal{R}(A)}^\perp = \mathcal{Y}$, and therefore the operator $\tilde{A}$ is *densely defined*, and it is defined on all of $\mathcal{Y}$ if and only if $\mathcal{R}(A)$ is closed.

We will not prove the following theorem, but it is interesting:

Theorem 2.8. *The Moore-Penrose inverse $A^\dagger$ is continuous if and only if $\mathcal{R}(A)$ is closed.*

The following characterises all important facts about the Moore-Penrose inverse:

Theorem 2.9 (Moore-Penrose equations)**.** *The operator $A^\dagger$ satisfies the following equations:*

(1) $A^\dagger A = P_{\mathcal{N}(A)^\perp}$;

(2) $AA^\dagger = P_{\overline{\mathcal{R}(A)}}\restriction_{\mathcal{D}(A^\dagger)}$;

(3) $AA^\dagger A = A$;

(4) $A^\dagger A A^\dagger = A^\dagger$.

Conversely, if any linear operator $B \colon \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp \to \mathcal{N}(A)^\perp$ satisfies $BA = P_{\mathcal{N}(A)^\perp}$ and $AB = P_{\overline{\mathcal{R}(A)}}\restriction_{\mathcal{D}(A^\dagger)}$ then $B = A^\dagger$.

Proof. We have

$$A^\dagger A u = \tilde{A}^{-1} A P_{\mathcal{N}(A)^\perp} u = P_{\mathcal{N}(A)^\perp} u,$$

which proves (1). Furthermore, we have

$$AA^\dagger f = A\tilde{A}^{-1} P_{\overline{\mathcal{R}(A)}} f = P_{\overline{\mathcal{R}(A)}} f,$$

which proves (2). Finally, (3) follows from (1) and (4) follows from (2).

Now, suppose B satisfies (1) and (2). First we show that $B\restriction_{\mathcal{R}(A)} = \tilde{A}^{-1}$, then we show that $B\restriction_{\mathcal{R}(A)^\perp} = 0$. This shows that $B = A^\dagger$. Let $f = Au \in \mathcal{R}(A)$ with $u \in \mathcal{N}(A)^\perp$, then

$$Bf = BAu = P_{\mathcal{N}(A)^\perp} u = u = \tilde{A}^{-1} f, \quad \text{so } B\restriction_{\mathcal{R}(A)} = \tilde{A}^{-1}.$$

Finally, let $f \in \mathcal{R}(A)^\perp$, then $ABf = P_{\overline{\mathcal{R}(A)}} f = 0$, and since $Bf \in \mathcal{N}(A)^\perp$ this implies $Bf = 0$. We conclude that $B\restriction_{\mathcal{R}(A)^\perp} = 0$, and this concludes the proof. □

The Moore-Penrose inverse has the important property that it maps every f in its domain to the corresponding minimum-norm least-squares solution:

Theorem 2.10. *For every $f \in \mathcal{D}(A^\dagger)$, the minimum-norm solution $u^\dagger$ to eq. (1) is given by $u^\dagger = A^\dagger f$.*

Proof. Since $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$, we know that there exists a unique minimum-norm solution $u^\dagger \in \mathcal{N}(A)^\dagger$. We write

$$u^\dagger = P_{\mathcal{N}(A)^\perp}(u^\dagger) = A^\dagger A u^\dagger = A^\dagger P_{\overline{\mathcal{R}(A)}} f = A^\dagger A A^\dagger f = A^\dagger f.$$

□

Remark. We can also consider the normal equation $A^* f = A^* A u$ as a least-squares problem, whose minimum-norm solution is $(A^*A)^\dagger A^* f$. It is clear that this expression must equal the minimum-norm solution $u^\dagger$ from eq. (1).

2.2 Compact operators

Definition 2.11. Let $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then A is called *compact* if for any bounded $B \subseteq \mathcal{X}$, the image $A(B)$ is precompact in $\mathcal{Y}$. The set of compact operators in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is denoted $\mathcal{K}(\mathcal{X}, \mathcal{Y})$.

Lemma 2.12. *Let $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then A is compact if and only if, for every bounded sequence $(x_n) \subseteq X$, the sequence $(Ax_n) \subseteq Y$ has a convergent subsequence.*

Theorem 2.13. *Let $A \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ with $\dim(\mathcal{R}(A)) = \infty$. Then $A^\dagger$ is discontinuous.*

Proof. If $\dim \mathcal{R}(A) = \infty$, then $\mathcal{X}$ and $\mathcal{N}(A)^\perp$ are infinite-dimensional as well. Chose an orthonormal sequence $(x_n) \subseteq \mathcal{N}(A)^\perp$, then after taking a subsequence if necessary, we can assume that $f_n \coloneqq Ax_n$ converges. However, we have

$$\left\|A^\dagger(f_n - f_m)\right\|^2 = \left\|A^\dagger A(x_n - x_m)\right\|^2 = \left\|P_{\mathcal{N}(A)^\perp}(x_n - x_m)\right\|^2 = \|x_n - x_m\|^2 = 2,$$

and in particular the sequence $(A^\dagger f_n)$ does not converge. This shows that $A^\dagger$ is discontinuous. □

In particular, combining this with theorem 2.8 shows that the range of a compact operator is always open, and that not every element in $\mathcal{Y}$ has a least-squares solution.

We will need the following theorem, an infinite-dimensional analogue of the spectral theorem:

Theorem 2.14 (Eigenvalue decomposition of self-adjoint compact operators)**.** *Let $\mathcal{X}$ be a Hilbert space, and $A \in \mathcal{K}(\mathcal{X}, \mathcal{X})$ self-adjoint. Then there exists an orthonormal basis (x_j) of $\overline{\mathcal{R}(A)}$ and a sequence of eigenvalues $|\lambda_1| \geq |\lambda_2| \geq \cdots > 0$ such that for all $u \in \mathcal{X}$ we have*

$$Au = \sum_{j=1}^{\infty} \lambda_j \langle u, x_j \rangle x_j.$$

The sequence (λ_j) is either finite or converges to 0.

The previous theorem gives rise to an infinite-dimensional analogue of the SVD:

Theorem 2.15. *Let $A \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$. Then there exists a (not necessarily infinite) sequence $\sigma_1 \geq \sigma_2 \geq \cdots > 0$ converging to 0, and orthonormal bases (x_j), (y_j) of $\mathcal{N}(A)^\perp$ and $\overline{\mathcal{R}(A)}$ respectively, such that*

$$Ax_j = \sigma_j y_j, \quad A^* y_j = \sigma_j x_j \quad \text{for all } j \in \mathbb{N},$$

and such that for all $u \in \mathcal{X}$ and $f \in \mathcal{Y}$ we have

$$Au = \sum_{j=1}^{\infty} \sigma_j \langle u, x_j \rangle y_j, \quad A^* f = \sum_{j=1}^{\infty} \sigma_j \langle f, y_j \rangle x_j.$$

The sequence $\{(\sigma_j, x_j, y_j)\}$ is called the singular value decomposition (SVD) of A.

Proof. Define $B \coloneqq A^*A$ and $C \coloneqq AA^*$, which are both compact, self-adjoint, and positive semi-definite operators. By the previous theorem, we can write

$$Cf = \sum_{j=1}^{\infty} \sigma_j^2 \langle f, y_j \rangle y_j,$$

where (y_j) is a basis of $\overline{\mathcal{R}(AA^*)} = \overline{\mathcal{R}(A)}$ and (σ_j) is a positive decreasing sequence converging to 0.

Note that

$$BA^*y_j = A^*AAy_j = A^*Cy_j = A^*\sigma^2 y_j = \sigma_j^2 A^* y_j,$$

so A^*y_j is an eigenvector of B with eigenvector σ_j^2.

We show that $\left(\frac{A^*y_j}{\sigma_j}\right)$ is an orthonormal basis of $\mathcal{R}(A)^\perp$. is an orthonormal basis of $\mathcal{N}(A)^\perp$: their inner product is given by

$$\left\langle \frac{A^*y_j}{\sigma_j}, \frac{A^*y_k}{\sigma_k} \right\rangle = \frac{1}{\sigma_j \sigma_k} \langle y_j, Cy_k \rangle = \frac{\sigma_k}{\sigma_j} \langle y_j, y_k \rangle = 0,$$

and since the (y_j) are a basis of $\overline{\mathcal{R}(A)} = \mathcal{N}(A^*)^\perp$ it is clear that the span of (A^*y_j) is dense in $\overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp$.

If we choose $x_j = \frac{A^*y_j}{\sigma_j}$, we find by construction that $A^*y_j = \sigma_j x_j$ and

$$Ax_j = \frac{AA^*y_j}{\sigma_j} = \frac{Cy_j}{\sigma_j} = \sigma_j y_j.$$

Finally, we see that

$$Au = \sum_{j=1}^{\infty} \langle u, x_j \rangle Ax_j = \sum_{j=1}^{\infty} \sigma_j \langle u, x_j \rangle y_j \quad \text{and} \quad A^*f = \sum_{j=1}^{\infty} \langle f, y_j \rangle A^* y_j = \sum_{j=1}^{\infty} \sigma_j \langle f, y_j \rangle x_j.$$

□

Theorem 2.16. *Let $A \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ with SVD $\{(\sigma_j, x_j, y_j)\}$ and let $f \in \mathcal{D}(A^\dagger)$. Then*

$$A^\dagger f = \sum_{j=1}^{\infty} \frac{1}{\sigma_j} \langle f, y_j \rangle x_j.$$

Remark. Note that this is comparable to $A^*f = \sum_{j=1}^{\infty} \sigma_j \langle f, y_j \rangle x_j$, except that A^* is a smoothing operator (since $\sigma_j \to 0$), while $A^\dagger$ does the opposite. Furthermore, $A^\dagger$ amplifies the right singular vectors corresponding to small singular values the most — intuitively, the corresponding left singular vectors are vectors where A doesn't "see much".

Proof. Define $Bf = \sum_{j=1}^{\infty} \frac{1}{\sigma_j} \langle f, y_j \rangle x_j$. Then by theorem 2.9, we must check that $BA = P_{\mathcal{N}(A)^\perp}$ and $AB = P_{\overline{\mathcal{R}(A)}} \restriction_{\mathcal{D}(A^\perp)}$.

For the first equation, we compute

$$BAu = \sum_{j=1}^{\infty} \frac{1}{\sigma_j} \left\langle \sum_{i=1}^{\infty} \sigma_i \langle u, x_i \rangle y_i, y_j \right\rangle x_j = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \frac{\sigma_i}{\sigma_j} \langle u, x_i \rangle \langle y_i, y_j \rangle x_j = \sum_{j=1}^{\infty} \langle u, x_j \rangle x_j.$$

Since (x_j) is a basis of $\mathcal{N}(A)^\perp$, this proves that $BA = P_{\mathcal{N}(A)^\perp}$.

For the second equation, an analogous computation gives $ABf = \sum_{i=1}^{\infty} \langle f, y_i \rangle y_i$, and since (y_i) is a basis of $\overline{\mathcal{R}(A)}$, this proves that $AB = P_{\overline{\mathcal{R}(A)}} \restriction_{\mathcal{D}(A^\dagger)}$. □

Definition 2.17. Let $A \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ have SVD $\{(\sigma_j, x_j, y_j)\}$. We say that $f \in \mathcal{Y}$ satisfies the *Picard criterion* if

$$\sum_j \frac{|\langle f, y_j\rangle|^2}{\sigma_j^2} < \infty.$$

Note that the expression on the left corresponds to $\left\|A^\dagger f\right\|^2$ if $f \in \mathcal{D}(A^\dagger)$.

Theorem 2.18. *Let $f \in \overline{\mathcal{R}(A)}$. Then $f \in \mathcal{R}(A)$ if and only if f satisfies the Picard criterion.*

Proof. ' $\implies$ ' Write $f = Au$, then

$$\sum_j \frac{|\langle f, y_j\rangle|^2}{\sigma_j^2} = \sum_j \frac{|\langle Au, y_j\rangle|^2}{\sigma_j^2} = \sum_j \frac{|\langle u, A^* y_j\rangle|^2}{\sigma_j^2} = \sum_j |\langle u, x_j\rangle|^2 < \infty.$$

' $\impliedby$ ' Define $u \coloneqq \sum_{j=1}^\infty \frac{1}{\sigma_j}\langle f, y_j\rangle x_j$ (note that by assumption this sum converges). Then

$$Au = A\sum_{j=1}^\infty \frac{1}{\sigma_j}\langle f, y_j\rangle x_j = \sum_{j=1}^\infty \frac{1}{\sigma_j}\langle f, y_j\rangle A x_j = \sum_{j=1}^\infty \langle f, y_j\rangle y_j = P_{\overline{\mathcal{R}(A)}} f = f,$$

so $Au = f$ which implies $f \in \mathcal{R}(A)$. □

We have seen that the stability of $A^\dagger$ depends on the speed of decay of the singular values (σ_j). We formalise this:

Definition 2.19. Let $A \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ have singular values (σ_j). Then the ill-posed inverse problem $Au = f$ is called *mildly ill-posed* if the σ_j decay polynomially (i.e., $\frac{1}{\sigma_n} \leq Cn^\gamma$ for some C, γ) and *severely ill-posed* otherwise.

Example 2.20. Consider the heat equation with initial conditions and boundary values:

$$\begin{cases} v_t - v_{xx} = 0 & (x,t) \in (0,\pi) \times \mathbb{R}_{>0}, \\ v(0,t) = v(\pi,t) = 0 & t \geq 0, \\ v(x,0) = u(x) & x \in (0,\pi), \\ v(x,T) = f(x) & x \in (0,\pi). \end{cases}$$

Then the forward problem is to determine f given u, while the inverse problem is to determine u given f. The solution for the foward problem is given by

$$f = Au \coloneqq \sum_{j=1}^\infty e^{-j^2 T}\langle u, \sin(jx)\rangle \sin(jx),$$

and the eigenvalues are therefore $\sigma_j = e^{-j^2 T}$. Since these clearly decay exponentially, this problem is severely ill-posed.

3 Classical regularisation theory

Let $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ such that $\mathcal{R}(A)$ is not closed (this happens for example when A is compact and does not have finite rank), and consider the inverse problem $Au = f$. Suppose we measure not f, but noisy data f_δ such that $\|f_\delta - f\| \leq \delta$. Then since $A^\dagger$ is discontinuous, we cannot expect that $A^\dagger f_\delta \to A^\dagger f$ as $\delta \to 0$. Therefore, we must replace $A^\dagger$ by operators that approximate it.

Definition 3.1. Let $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. A family $(R_\alpha)_{\alpha>0}$ of continuous operators is called a *regularisation* of $A^\dagger$ if

$$\lim_{\alpha \to 0} R_\alpha f = A^\dagger f \quad \text{for all } f \in \mathcal{D}(A^\dagger).$$

If all R_α are linear (TODO: and bounded?), then we speak of a *linear regularisation* of $A^\dagger$.

Theorem 3.2 (Banach-Steinhaus)**.** *Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\{A_\alpha\} \subseteq \mathcal{B}(\mathcal{X}, \mathcal{Y})$ a family of pointwise bounded operators. Then $\{A_\alpha\}$ is bounded in norm.*

Corollary 3.3. *Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $(A_j) \subseteq \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then (A_j) converges pointwise to some $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ if and only if $\{A_j\}$ is norm-bounded and converges pointwise on some dense subset $\mathcal{X}' \subseteq \mathcal{X}$.*

Theorem 3.4. *Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and $(R_\alpha)_{\alpha>0}$ a linear regularisation. If $A^\dagger$ is not continuous, (R_α) is not norm-bounded. In particular, there exists $f \in \mathcal{Y}$ with $\|R_\alpha f\| \to \infty$.*

Proof. Suppose (R_α) is norm-bounded. Let $\alpha_j \to 0$, then we know that $R_{\alpha_j} \to A^\dagger$ pointwise on $\mathcal{D}(A^\dagger)$. Since $\mathcal{D}(A^\dagger)$ is dense in $\mathcal{Y}$, corollary 3.3 then tells us that $A^\dagger$ is bounded and therefore continuous, a contradiction.

By the Banach-Steinhaus theorem, if (R_α) is not norm-bounded, it is not pointwise bounded, so there must exist $f \in \mathcal{Y}$ such that $\{\|R_\alpha f\|\}$ is not bounded. □

Recap 3.5. Recall that any bounded sequence in a Hilbert space has a weakly convergent subsequence.

Theorem 3.6. *Let $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and (R_α) a linear regularisation of $A^\dagger$. If $\{\|AR_\alpha\|\}_{\alpha>0}$ is bounded, then $\|R_\alpha f\| \to \infty$ as $\alpha \to 0$ for every $f \notin \mathcal{D}(A^\dagger)$.*

Proof. Define $u_\alpha := R_\alpha f$ for $f \notin \mathcal{D}(A^\dagger)$, and assume there exists a sequence $\alpha_k \to 0$ such that $\{\|u_{\alpha_k}\|\}$ is bounded. After taking a subsequence if necessary, we may assume that $u_{\alpha_k} \rightharpoonup u$ for some $u \in \mathcal{X}$, and therefore we also have $Au_{\alpha_k} \rightharpoonup Au$.

We also have $\lim_{\alpha\to 0} AR_\alpha f = AA^\dagger f = P_{\overline{\mathcal{R}(A)}} f$ for $f \in \mathcal{D}(A^\dagger)$, and since we assumed $\{AR_\alpha\}$ was norm-bounded, by corollary 3.3 we have $\lim_{\alpha\to 0} AR_\alpha f = P_{\overline{\mathcal{R}(A)}} f$ for all $f \in \mathcal{Y}$.

Since Au_{α_k} is convergent and has weak limit Au, it must also have limit Au, so we find $Au = P_{\overline{\mathcal{R}(A)}} f$ so $f \in \mathcal{D}(A^\dagger)$, a contradiction. □

We need some process to choose a parameter. To this end, note that we have

$$\left\|R_\alpha f_\delta - A^\dagger f\right\| \leq \|R_\alpha(f_\delta - f)\| + \left\|(R_\alpha - A^\dagger)f\right\| \leq \delta\|R_\alpha\| + \left\|(R_\alpha - A^\dagger)f\right\|. \tag{3}$$

The first term is called the *data error* and is unbounded for $\alpha \to 0$, and the second term is called the *approximation error* which does vanish for $\alpha \to 0$. Therefore, we want to choose α small enough to have a low approximation error, while keeping the data error at bay.

3.1 Parameter choice rules

Definition 3.7. A function $\alpha\colon \mathbb{R}_{>0} \times \mathcal{Y} \to \mathbb{R}_{>0}\colon (\delta, f_\delta) \mapsto \alpha(\delta, f_\delta)$ is called a *parameter choice rule* (PCR). We distinguish three types:

1. An *a priori* PCR depends only on δ;
2. An *a posteriori* PCR depends on both δ and f_δ;
3. A *heuristic* PCR depends only on f_δ.

Definition 3.8. Let $(R_\alpha)_{\alpha>0}$ be a regularisation of $A^\dagger$ and α a parameter choice rule. We call (R_α, α) a *convergent regularisation* if

$$\lim_{\delta\to 0} \sup_{f_\delta : \|f - f_\delta\| \le \delta} \left\| R_\alpha f_\delta - A^\dagger f \right\| = 0$$

and

$$\lim_{\delta\to 0} \sup_{f_\delta : \|f - f_\delta\| \le \delta} \alpha(\delta, f_\delta) = 0. \tag{4}$$

3.1.1 A priori parameter choice rules

We will not prove the following theorem, which guarantees the existence of a priori PCRs:

Theorem 3.9. *Let $(R_\alpha)_{\alpha>0}$ be a regularisation of $A^\dagger$. Then there exists an a priori PCR $\alpha = \alpha(\delta)$ such that (R_α, α) is convergent.*

We can characterise PCRs in the following way:

Theorem 3.10. *Let $(R_\alpha)_{\alpha>0}$ be a linear regularisation of $A^\dagger$, and $\alpha = \alpha(\delta)$ an a priori PCR. Then (R_α, α) is convergent if and only if*

$$\lim_{\delta\to 0} \delta \left\| R_{\alpha(\delta)} \right\| = 0 \quad \textit{and} \quad \lim_{\delta\to 0} \alpha(\delta) = 0.$$

Proof. " $\implies$ " Suppose (R_α, α) is convergent. It is clear that $\lim_{\delta\to 0} \alpha(\delta) = 0$ by eq. (4). Suppose $\lim_{\delta\to 0} \delta \left\| R_{\alpha(\delta)} \right\| \neq 0$. Then there exists a sequence $(\delta_k) \to 0$ and a constant $C > 0$ such that $\delta_k \left\| R_{\alpha(\delta_k)} \right\| \ge C$ for all k. This implies we can find a sequence $(g_k) \subseteq \mathcal{Y}$ with $\|g_k\| = 1$ and $\delta_k \left\| R_{\alpha(\delta_k)} g_k \right\| \ge C$ for all k.

Now let $f \in \mathcal{D}(A^\dagger)$ and define $f_k := f + \delta_k g_k$, then clearly we have $f_k \to f$, but also

$$C \le \left\| R_{\alpha(\delta_k)}(\delta_k g_k) \right\| = \left\| R_{\alpha(\delta_k)}(f_{\delta_k} - f) \right\| \le \left\| R_{\alpha(\delta_k)} f_{\delta_k} - A^\dagger f \right\| + \left\| (R_{\alpha(\delta_k)} - A^\dagger) f \right\|.$$

In particular we find that $\left\| (R_{\alpha(\delta_k)} - A^\dagger) f \right\| \ge C$, so clearly R_α is not convergent.

" $\impliedby$ " This follows immediately from eq. (3). □

A problem with a priori PCRs is that they are scale-invariant: if $\alpha = \alpha(\delta)$ gives a convergent regularisation, then $\hat{\alpha} = \alpha(k\delta)$ also gives a convergent regularisation for any k. In practice, it is not always clear which scale should be chosen.

3.1.2 A posteriori parameter choice rules

Let $f \in \mathcal{D}(A^\dagger)$ and f_δ s.t. $\|f - f_\delta\| \le \delta$. Letting $u^\dagger$ denote the minimum-norm solution of the problem $Au = f$, and defining $\mu := \left\| Au^\dagger - f \right\| = \inf_{u\in\mathcal{X}} \|Au - f\|$, we see that

$$\left\| Au^\dagger - f_\delta \right\| \le \left\| Au^\dagger - f \right\| + \|f - f_\delta\| \le \mu + \delta.$$

Therefore, it is not useful to choose $\alpha(\delta, f_\delta)$ with $\|Au_\alpha - f_\delta\| < \mu + \delta$: if this is the case, we are most likely overfitting.

This motivates *Morozov's discrepancy principle*:

Definition 3.11. Let (R_α) be a (TODO: linear?) regularisation of $A^\dagger$ and assume $\mathcal{R}(A)$ is dense in $\mathcal{Y}$. Fix $\eta > 1$, and define

$$\alpha(\delta, f_\delta) = \sup\{\alpha > 0 : \|AR_\alpha f_\delta - f_\delta\| \leq \eta\delta\}.$$

Then $\alpha(\delta, f_\delta)$ is said to satisfy *Morozov's discrepancy principle*.

It can be shown that the above α indeed gives a convergent regularisation.

3.1.3 Heuristic parameter choice rules

Heuristic parameter choice rules unfortunately only work if the original problem was well-posed:

Theorem 3.12 (Bakushinskii)**.** *Let (R_α) be a regularisation of $A^\dagger$ and suppose there exists a heuristic parameter choice rule α such that (R_α, α) is convergent. Then $A^\dagger$ is continuous from $\mathcal{Y}$ to $\mathcal{X}$.*

3.2 Spectral regularisation

We will now start with specific examples of regularisations. Spectral regularisations are derived from the spectral decomposition

$$A^\dagger f = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, y_j \rangle x_j.$$

We construct a regularisation by replacing σ_j^{-1} by some function $g_\alpha(\sigma_j)$, i.e.,

$$R_\alpha f = \sum_{j=1}^{\infty} g_\alpha(\sigma_j) \langle f, y_j \rangle x_j. \tag{5}$$

Let us explore which conditions g_α must satisfy:

Theorem 3.13. *Let, for $\alpha > 0$, the function $g_\alpha \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ satisfy*

1. $\lim_{\alpha \to 0} g_\alpha(\sigma) = \frac{1}{\sigma}$ *for all* $\sigma > 0$;
2. $g_\alpha(\sigma) \leq C_\alpha$ *for some* $C_\alpha > 0$;
3. $\sup_{\alpha,\sigma} \sigma g_\alpha(\sigma) \leq \gamma$ *for some* $\gamma > 0$.

Then collection (R_α) defined by eq. (5) is a linear regularisation of $A^\dagger$, and in particular, we have $\|R_\alpha\| \leq C_\alpha$.

Proof. From condition 2 it follows that all R_α are bounded. Since

$$\langle f, y_j \rangle = \langle P_{\overline{\mathcal{R}(A)}} f, y_j \rangle = \langle AA^\dagger f, y_j \rangle = \langle A^\dagger f, A^* y_j \rangle = \sigma_j \langle u^\dagger, x_j \rangle,$$

we compute

$$(R_\alpha - A^\dagger) f = \sum_j (g_\alpha(\sigma_j) - \sigma_j)^{-1} \langle f, y_j \rangle x_j = \sum_j (\sigma_j g_\alpha(\sigma_j) - 1) \langle u^\dagger, x_j \rangle x_j,$$

and since $\sigma g_\alpha(\sigma_j) \leq \gamma$, we have $(\sigma_j g_\alpha(\sigma_j) - 1)^2 \leq 1 + \gamma^2$, so that

$$\left\|(R_\alpha - A^\dagger) f\right\|^2 \leq (1 + \gamma^2) \left\|u^\dagger\right\|^2 < \infty.$$

Since $\left\|(R_\alpha - A^\dagger) f\right\|$ is finite, we may apply the reverse Fatou lemma to the sum and obtain

$$\limsup_{\alpha \to 0} \left\|(R_\alpha - A^\dagger) f\right\|^2 \leq \sum_j \left(\sigma_j \limsup_{\alpha \to 0} g_\alpha(\sigma_j) - 1 \right)^2 \langle u^\dagger, x_j \rangle^2 = 0,$$

and therefore $R_\alpha f \to A^\dagger f$ as $\alpha \to 0$. □

Example 3.14. The first, very simple example is the *truncated SVD*: we simply define

$$g_\alpha(\sigma) = \begin{cases} 1/\sigma & \sigma \geq \alpha, \\ 0 & \sigma < \alpha. \end{cases}$$

It is easy to check that g_α satisfies the conditions of theorem 3.13, and that all R_α are finite-rank operators with $\|R_\alpha\| \leq \frac{1}{\alpha}$. Therefore, if we choose $\alpha = \alpha(\delta)$ such that $\delta/\alpha(\delta) \to 0$, then we obtain a convergent regularisation.

This also highlights the problem with this method: as δ gets smaller, we need more and more singular vectors which are generally expensive to compute.

Example 3.15. The second example is *Tikhonov regularisation*. Here, we define $g_\alpha(\sigma) = \frac{\sigma}{\sigma^2+\alpha}$, and again it is easily checked that the conditions of theorem 3.13 are satisfied, noting that

$$\frac{\sigma}{\sigma^2 + \alpha} \leq \frac{\sigma}{2\sigma\sqrt{\alpha}} = \frac{1}{2\sqrt{\alpha}} =: C_\alpha.$$

Therefore, if $\delta/\sqrt{\alpha(\delta)} \to 0$, the regularisation is convergent.

This method does not require computing the SVD of A: it is easily shown that $u_\alpha \coloneqq R_\alpha f$ is the unique solution to the *regularised normal equation*

$$(A^*A + \alpha I)u_\alpha = A^*f.$$

While $A^*A + \alpha I$ is always invertible, computing the inverse is expensive, so we usually use some approximation of the inverse.

Finally, it can also be shown that

$$u_\alpha = \min_{u \in \mathcal{X}} \|Au - f\|^2 + \alpha\|u\|^2,$$

so we can also view u_α as the solution of an optimisation problem.

4 Variational regularisation

4.1 Background

4.1.1 Banach spaces and weak convergence

A *Banach space* $\mathcal{X}$ is a complete normed vector space. We define the *dual space* $\mathcal{X}^* \coloneqq \mathcal{L}(X, \mathbb{R})$, and for $p \in \mathcal{X}^*, u \in \mathcal{X}$ we usually write $\langle p, u\rangle$ instead of $p(u)$. For any $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ we define the *adjoint* $A^*\colon \mathcal{Y}^* \to \mathcal{X}^*$ by $\langle A^* p, u\rangle \coloneqq \langle p, Au\rangle$ for all $p \in \mathcal{X}^*, u \in \mathcal{X}$. The dual space $\mathcal{X}'$ is equipped with the norm

$$\|p\|_{\mathcal{X}^*} \coloneqq \sup_{\|u\|\leq 1} \langle p, u\rangle,$$

and with this norm $\mathcal{X}^*$ is a Banach space.

The *bi-dual space* is defined as $\mathcal{X}^{**} \coloneqq (\mathcal{X}^*)^*$. The mapping $E\colon \mathcal{X} \to (\mathcal{X})^{**}$ defined by $\langle E(u), p\rangle \coloneqq \langle p, u\rangle$ is a continuous linear isometry, and we will regard $\mathcal{X}$ as a subspace of $\mathcal{X}^{**}$ using this isometry. If $\mathcal{X} = \mathcal{X}^{**}$ (i.e., E is surjective), the space $\mathcal{X}$ is called *reflexive*. A space $\mathcal{X}$ is called *separable* if $\mathcal{X}$ has a countable dense subset.

A sequence $(u_k) \subseteq \mathcal{X}$ is said to *converge weakly* to $u \in \mathcal{X}$, denoted $u_k \rightharpoonup u$, if $\langle p, u_k\rangle \to \langle p, u\rangle$ for all $p \in \mathcal{X}^*$.

A sequence $(p_k) \subseteq \mathcal{X}^*$ is said to *converge weakly-** to $p \in \mathcal{X}'$, denoted $p_k \overset{*}{\rightharpoonup} p$, if $\langle p_k, u\rangle \to \langle p, u\rangle$ for all $u \in \mathcal{X}$.

Theorem 4.1. *Let $\mathcal{X}$ be Banach, then the unit ball is compact in $\mathcal{X}^*$ w.r.t. the weak-$*$ topology. If $\mathcal{X}$ is separable, then the weak-$*$ topology is metrisable and every bounded sequence in $\mathcal{X}^*$ has a weakly-$*$ convergent subsequence.*

Theorem 4.2. *Let $\mathcal{X}$ be reflexive, then every bounded sequence in $\mathcal{X}$ has a weakly convergent subsequence.*

We define $\overline{\mathbb{R}} \coloneqq \mathbb{R} \cup \{\pm\infty\}$.

Definition 4.3. Let $\mathcal{X}$ be a Banach space with topology τ_X. A functional $E\colon \mathcal{X} \to \overline{\mathbb{R}}$ is said to be *sequentially lower-semicontinuous with respect to* $\tau_{\mathcal{X}}$ or simply $\tau_{\mathcal{X}}$-LSC if

$$E(u) \leq \liminf_{n\to\infty} E(u_n) \quad \text{if } u_n \overset{\tau}{\to} u.$$

Specifically, if $\tau_{\mathcal{X}}$ is the weak topology, then E is called *weakly* LSC. If $\tau_{\mathcal{X}}$ is the topology induced by the norm on $\mathcal{X}$, then E is called *strongly* LSC or simply LSC.

Lemma 4.4. *Let $E\colon \mathcal{X} \to \overline{\mathbb{R}}$ be $\tau_{\mathcal{X}}$-LSC and $a \in \mathbb{R}$. Then $\{u \in \mathcal{X} \mid E(u) \leq a\}$ is $\tau_{\mathcal{X}}$-closed.*

Proof. Let $(u_n) \overset{\tau}{\to} u$ and suppose $E(u_n) \leq a$ for all n. Then clearly $E(u) \leq \liminf E(u_n) \leq a$. □

4.1.2 Convex analysis

Definition 4.5. Let $C \subseteq \mathcal{X}$. Then the *characteristic function* of C is defined as

$$\chi_C(u) \coloneqq \begin{cases} 0, & u \in C, \\ \infty, & u \notin C. \end{cases}$$

Using characteristic functions, we have $\min_{u\in C} E(u) = \min_{u\in\mathcal{X}} E(u) + \chi_C(u)$.

Definition 4.6. Let $E\colon \mathcal{X} \to \overline{\mathbb{R}}$, then the *effective domain* is $\operatorname{dom}(E) \coloneqq \{u \mid E(u) < \infty\}$.

The functional E is called *proper* if $\operatorname{dom}(E) \neq \varnothing$.

Definition 4.7. A functional $E\colon \mathcal{X} \to \overline{\mathbb{R}}$ is called:

1. *convex* if for all $u \neq v \in \mathcal{X}$ and $\lambda \in (0,1)$ we have $E(\lambda u + (1-\lambda)v) \leq \lambda E(u) + (1-\lambda)E(v)$;
2. *strictly convex* if the above inequality is strict;
3. *strongly convex* with constant $\vartheta > 0$ if $u \mapsto E(u) - \vartheta\|u\|^2$ is convex.

Note that $C \subseteq \mathcal{X}$ is a convex set if and only if χ_C is a convex function.

Lemma 4.8. *Nonnegative linear combinations of convex functionals are convex. If one of the components is strictly convex, then the nonnegative linear combination is also strictly convex.*

Definition 4.9. Let $E\colon \mathcal{X} \to \overline{\mathbb{R}}$ be a functional. We define the *Fenchel conjugate*

$$E^*\colon \mathcal{X}^* \to \overline{\mathbb{R}}\colon p \mapsto \sup_{u\in\mathcal{X}} [\langle p, u\rangle - E(u)].$$

Theorem 4.10. *For any $E\colon \mathcal{X} \to \mathbb{R}$ we have $E^{**}\restriction_{\mathcal{X}} \leq E$. If E is proper and LSC, then $E^{**}\restriction_{\mathcal{X}} = E$.*

Definition 4.11. A functional $E\colon \mathcal{X} \to \overline{\mathbb{R}}$ is called *subdifferentiable* at $u \in \mathcal{X}$ if there exists a $p \in \mathcal{X}^*$ such that

$$E(v) \geq E(u) + \langle p, v-u\rangle \quad \text{for all } v \in \mathcal{X}.$$

In this case, we call p a *subgradient* of E at position u. The collection of all subgradients of E at u is denoted by $\partial E(u)$ and is called the *subdifferential* of E at u.

Lemma 4.12. *Let $E\colon \mathcal{X} \to \overline{\mathbb{R}}$ be convex, then E is subdifferentiable at all points $u \in \operatorname{dom}(E)$ where E is continuous. If E is also proper, then E is not subdifferentiable at any $u \notin \operatorname{dom}(E)$.*

Theorem 4.13. *Let $E\colon \mathcal{X} \to \overline{\mathbb{R}}$ be proper and convex and $u \in \operatorname{dom}(E)$. Then $\partial E(u)$ is convex and weakly-$*$ compact in $\mathcal{X}^*$.*

Theorem 4.14. *Let E, F be proper LSC convex functionals and $u \in \operatorname{dom}(E) \cap \operatorname{dom}(F)$ such that at least one of E and F is continuous at u. Then $\partial(E+F)(u) = \partial E(u) + \partial F(u)$.*

Theorem 4.15. *Let E be convex. Then u is a global minimiser of E if and only if $0 \in \partial E(u)$.*

Definition 4.16. Let E be convex, $u, v \in \mathcal{X}$, $E(v) < \infty$ and $q \in \partial E(v)$. Then the *Bregman distance* of E between u and v is defined as

$$D_E^q(u,v) \coloneqq E(u) - E(v) - \langle q, u-v\rangle \geq 0.$$

If we also have $E(u) < \infty, p \in \partial E(u)$, then we define the *symmetric Bregman distance*

$$D_E^{p,q}(u,v) \coloneqq D_E^p(v,u) + D_E^q(u,v) = \langle p-q, u-v\rangle.$$

Definition 4.17. For $p > 0$, a functional E is called *absolutely p-homogeneous* if $E(\lambda u) = |\lambda|^p E(u)$ for all $\lambda \in \mathbb{R}, u \in \mathcal{X}$.

Proposition 4.18. *Let E be a convex, proper and absolutely one-homogeneous, and $p \in \partial E(u)$. Then:*

1. *$E(u) = \langle p, u\rangle$;*
2. *$D^p(v,u) = E(v) - \langle p, v\rangle$ for all $v \in \mathcal{X}$;*
3. *$E^*(p) = \chi_{\partial E(0)}(p)$.*

Furthermore, we have the following:

Proposition 4.19. *Let E be proper, convex, and absolutely one-homogeneous, and let $u \in \mathcal{X}$. Then $p \in \partial E(u)$ if and only if $p \in \partial E(0)$ and $\langle p, u\rangle = E(u)$.*

4.1.3 Minimisers

Definition 4.20. We say that $u^* \in \mathcal{X}$ is a *minimiser* of a functional E if u minimises E and $E(u) < \infty$.

Definition 4.21. A functional E is called *coercive* if $\|u_j\| \to \infty \implies |E(u_j)| \to \infty$.

Lemma 4.22. *Let E be proper, coercive and bounded from below. Then $\inf_{u\in\mathcal{X}} E(u) > -\infty$ and there exists a (bounded) minimising sequence (u_j) with $E(u_j) \to \inf_u E(u)$.*

Theorem 4.23 (Direct method)**.** *Let $\mathcal{X}$ be Banach and $\tau_\mathcal{X}$ a topology on $\mathcal{X}$ such that any bounded sequence in $\mathcal{X}$ has a $\tau_\mathcal{X}$ convergent subsequence. Then any proper, bounded from below, coercive, τ_X-LSC functional has a minimiser.*

Proof. Since E is bounded from below, we have $\inf_u E(u) > -\infty$, so there exists a bounded minimising sequence (u_j), which we can assume is $\tau_\mathcal{X}$ convergent with limit u^* after taking a subsequence if necessary. By lower-semicontinuity of E we have

$$E(u^*) \le \liminf_{k\to\infty} E(u_j) = \lim_{j\to\infty} E(u_j) = \inf_u E(u),$$

so u^* is a minimiser. □

Theorem 4.24. *If a strictly convex functional has a minimiser, it is unique.*

Proof. Suppose $u \neq v$ are two minimisers, then by strict convexity, we have $E(\frac{1}{2}u + \frac{1}{2}v) < E(u)$, a contradiction. □

4.1.4 Duality in convex optimisation

Consider the *primal* optimisation problem

$$(P) \coloneqq \inf_{u\in\mathcal{X}} E(Au) + F(u),$$

where E, F are proper, convex and LSC, and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Since E is convex and LSC, we have $E = E^{**}$ so we can rewrite the primal problem as the *saddle point problem*

$$\inf_{u\in\mathcal{X}} \sup_{\eta\in\mathcal{Y}^*} \langle \eta, Au\rangle - E^*(\eta) + F(u).$$

Since $\inf\sup \ge \sup\inf$ always holds we have

$$(P) \ge \sup_{\eta\in\mathcal{Y}^*} \inf_{u\in\mathcal{X}} \langle \eta, y\rangle - E^*(\eta) + F(u) = \sup_{\eta\in\mathcal{Y}^*} -E^*(\eta) - F^*(-A^*\eta) =: (D).$$

The problem (D) is called the *dual problem*, and the fact that $(D) \le (P)$ is called *weak duality*. The value $(P) - (D)$ is called the *duality gap*, and if $(P) = (D)$, we speak of *strong duality*.

We have the following:

Theorem 4.25. *Suppose the function $E(Au) + F(u)$ is proper, convex, LSC and coercive. Suppose also that there exists $u_0 \in \mathcal{X}$ s.t. $F(u) < \infty$, $E(Au_0) < \infty$, and $E(y)$ is continuous at $y = Au_0$. Then:*

1. *The dual problem (D) has at least one solution $\hat{\eta}$;*
2. *There is no duality gap;*
3. *If (P) has an optimal solution $\hat{u}$, then we have*

$$A^*\hat{\eta} \in \partial F(\hat{u}), \quad -\hat{\eta} \in \partial E(A\hat{u}).$$

4.2 Regularisation properties

We consider the problem $Au = f$ where $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and where $\mathcal{Y}$ is a Banach space and $\mathcal{X}$ is the dual of a separable Banach space. Recall that in Tychonoff regularisation, we have

$$R_\alpha(f_\delta) = \arg\min_{u \in \mathcal{X}} \|Au - f_\delta\|^2 + \alpha\|u\|^2.$$

Here, the term $\|u\|^2$ is known as the *regularisation term*, and penalising large values of $\|u\|$ ensures regularity of the solution. Inspired by this, suppose we have any function $\mathcal{J}(u)$ as regulariser, then the *variational regularisation problem* is given by

$$R_\alpha f_\delta \in \arg\min_{u \in \mathcal{X}} \frac{1}{2}\|Au - f_\delta\|^2 + \alpha\mathcal{J}(u),$$

Definition 4.26. Let $u^\dagger_{\mathcal{J}}$ be a least-squares solution that minimises $\mathcal{J}$ over all least-squares solutions. Then we call $u^\dagger_{\mathcal{J}}$ a *$\mathcal{J}$-minimising solution* of the problem $Au = f$.

Convention. We will assume there exists t least one least-squares solution with a finite value of $\mathcal{J}$.

Lemma 4.27. *Let $\mathcal{J}(u) \coloneqq \sum_{i=1}^n \mathcal{J}_i(u)$, where each $\mathcal{J}_i(u)$ is convex and absolutely p_i-homogeneous ($p_i > 0$). Then $\mathcal{N}(J)$ is a linear subspace of $\mathcal{X}$.*

Proof. We note that $\mathcal{J}_i$ is nonnegative, since

$$0 = \mathcal{J}_i(0) = \mathcal{J}_i\left(\frac{1}{2}u - \frac{1}{2}u\right) \leq \frac{1}{2}\mathcal{J}_i(u) + \frac{1}{2}\mathcal{J}_i(-u) = \mathcal{J}_i(u).$$

Therefore we have $\mathcal{N}(\mathcal{J}) = \cap_i \mathcal{N}(\mathcal{J}_i)$, and we will show that every $\mathcal{N}(\mathcal{J}_i)$ is a subspace. Let $\lambda \in \mathbb{R}, u, v \in \mathcal{N}(\mathcal{J}_i)$, then

$$\mathcal{J}_i(\lambda u + v) = 2^{p_i}\mathcal{J}_i\left(\frac{\lambda u}{2} + \frac{v}{2}\right) \leq 2^{p_i - 1}(|\lambda|^{p_i}\mathcal{J}_i(u) + \mathcal{J}_i(v)) = 0.$$

so $\lambda u + v \in \mathcal{N}(\mathcal{J}_i)$. This completes the proof. □

Lemma 4.28. *Let the assumptions of lemma 4.27 be satisfied, and suppose $u \in \mathcal{X}, v \in \mathcal{N}(\mathcal{J})$. Then $\mathcal{J}(u + v) = \mathcal{J}(u)$.*

Proof. Clearly it suffices to prove this lemma for each $\mathcal{J}_i$. It is easily seen that for any $t \in (0, 1)$ we have

$$\mathcal{J}_i(u + v) = \mathcal{J}_i\left(t\frac{u}{t} + (1 - t)\frac{v}{1 - t}\right) \leq t\mathcal{J}_i\left(\frac{u}{t}\right) = t^{1-p_i}\mathcal{J}(u).$$

Letting $t \to 1$, we find $\mathcal{J}_i(u + v) \leq \mathcal{J}_i(u)$.

Similarly, we have

$$\mathcal{J}_i(u) = \mathcal{J}_i(u + v - v) = \mathcal{J}_i\left(t\frac{u + v}{t} + (1 - t)\frac{-v}{1 - t}\right) \leq t\mathcal{J}_i\left(\frac{u + v}{t}\right) = t^{1-p_i}\mathcal{J}_i(u + v),$$

and letting $t \to 1$ we obtain $\mathcal{J}_i(u) \leq \mathcal{J}_i(u + v)$, so $\mathcal{J}_i(u) = \mathcal{J}_i(u + v)$, so $\mathcal{J}(u) = \mathcal{J}(u + v)$. □

Recap 4.29. Let $\mathcal{X}$ be a Banach space and let $U \subseteq \mathcal{X}$ be a closed subspace of $\mathcal{X}$. Then U is called *complemented* in $\mathcal{X}$ if there exists a closed subspace $V \subseteq \mathcal{X}$ such that $\mathcal{X} = U \oplus V$.

In general, it is difficult to determine which closed subspaces of $\mathcal{X}$ are complemented. However, it is known that all **finite-dimensional** subspaces are complemented.

Lemma 4.30. *Suppose $\mathcal{J}$ is proper, convex, and satisfies the conditions of lemma 4.27. Suppose also that:*

(i) $\dim \mathcal{N}(\mathcal{J}) < \infty$ *(so that $\mathcal{X} = \mathcal{N}(\mathcal{J}) + \mathcal{X}_0$ for some closed subspace $\mathcal{X}_0$) and $\mathcal{J}$ is coercive on $\mathcal{X}_0$;*

(ii) $\mathcal{N}(A) \cap \mathcal{N}(\mathcal{J}) = \{0\}$.

Then the function $\Phi_\alpha(u) \coloneqq \frac{1}{2}\|Au - f\|^2 + \alpha\mathcal{J}(u)$ is coercive on $\mathcal{X}$ for any $\alpha > 0$.

Proof. Let $(u_j) \subseteq \mathcal{X}$ be a sequence such that $(\Phi_\alpha(u_j))$ is bounded. We will prove that (u_j) is bounded.

Decompose every u_j as $u_j = u_j^0 + u_j^{\mathcal{N}}$ with $u_j^0 \in \mathcal{X}_0$ and $u_j^{\mathcal{N}} \in \mathcal{N}(\mathcal{J})$. Now, since $(\Phi_\alpha(u_j))$ is bounded, the sequence $(\mathcal{J}(u_j)) = (\mathcal{J}(u_j^0))$ is bounded as well, and by coercivity of $\mathcal{J}$ on $\mathcal{X}^0$ we conclude that (u_j^0) is bounded.

Now define $\tilde{A} \coloneqq A\restriction_{\mathcal{N}(\mathcal{J})}$. Since $\mathcal{N}(A) \cap \mathcal{N}(\mathcal{J}) = \{0\}$, $\tilde{A}$ has an inverse on $A\mathcal{N}(\mathcal{J})$, and since $\dim A\mathcal{N}(\mathcal{J}) < \infty$, that inverse is bounded. Therefore we find

$$\|u_j^{\mathcal{N}}\| = \left\|\tilde{A}^{-1}(Au_j^{\mathcal{N}})\right\| \le C\left\|\tilde{A}u_j^{\mathcal{N}}\right\| = C\left\|\tilde{A}u_j - f - (Au_j^0 - f)\right\| \le C\big(\|Au_j - f\| + \|A\|\|u_j^0\| + \|f\|\big).$$

Since $\Phi_\alpha(u_j)$ is bounded, $\|Au_j - f\|$ is also bounded, and the other terms are also all bounded, so we conclude that $(u_j^{\mathcal{N}})$ is bounded, and therefore that (u_j) is bounded. □

Theorem 4.31. *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces with topologies $\tau_{\mathcal{X}}, \tau_{\mathcal{Y}}$. Assume that:*

(a) bounded sequences in $\mathcal{X}$ have $\tau_{\mathcal{X}}$-convergent subsequences;

(b) $\mathcal{J}\colon \mathcal{X} \to [0, \infty]$ is proper, convex, $\tau_{\mathcal{X}}$-LSC and satisfies the assumptions of lemma 4.30;

(c) A is continuous w.r.t. $\tau_{\mathcal{X}}$ and $\tau_{\mathcal{Y}}$;

(d) $\|\cdot\|_{\mathcal{Y}}$ is $\tau_{\mathcal{Y}}$-LSC.

Then:

1. *There exists a $\mathcal{J}$-minimising solution $u_{\mathcal{J}}^\dagger$ of the equation $Au = f$;*
2. *For any $\alpha > 0$, $f \in \mathcal{Y}$ there exists a minimiser*

$$u_\alpha = R_\alpha f \in \arg\min \frac{1}{2}\|Au - f\|^2 + \alpha\mathcal{J}(u).$$

Proof. 1. Let $\mathbb{L}$ be the set of least-squares solutions and $\mu \coloneqq \inf\{\|Au - f\| \mid u \in \mathcal{X}\}$, then we can write

$$\mathbb{L} = \{u \in \mathcal{X} : \|Au - f\| \le \mu\}.$$

Since $\|\cdot\|_{\mathcal{Y}}$ is $\tau_{\mathcal{Y}}$-LSC and $A\colon \tau_{\mathcal{X}} \to \tau_{\mathcal{Y}}$ is continuous, we have by lemma 4.4 that $\mathbb{L}$ is $\tau_{\mathcal{X}}$-closed. Now consider the problem

$$\inf_{u\in\mathbb{L}} \mathcal{J}(u) = \inf_{u\in\mathcal{X}} \mathcal{J}(u) + \chi_{\mathbb{L}}(u).$$

By assumption there exists $u \in \mathcal{L}$ with $\mathcal{J}(u) < \infty$, and the objective function $\mathcal{J} + \chi_{\mathcal{L}}$ is bounded from below (by 0). Using similar arguments as in the previous lemma (???), it can be shown that it is coercive. Finally, characteristic functions of closed sets are LSC, so $\chi_{\mathbb{L}}$ and $\mathcal{J}$ are both $\tau_{\mathcal{X}}$-LSC, and therefore their sum is $\tau_{\mathcal{X}}$-LSC as well. By the direct method (theorem 4.23), we conclude that a $\mathcal{J}$-minimising solution exists.

2. The objective function Φ_α is coercive by the previous lemma, and also bounded from below. It is easily seen that Φ_α is $\tau_{\mathcal{X}}$-LSC, and using the direct method we conclude that Φ_α has a minimiser.

□

Theorem 4.32. *Let the assumptions of theorem 4.31 be satisfied and suppose that* $\inf_{v\in\mathcal{X}} \|Av - f\| = 0$. *Let* $\alpha = \alpha(\delta)$ *satisfy*

$$\lim_{\delta\to 0} \alpha(\delta) = 0 = \limsup_{\delta\to 0} \frac{\delta^2}{\alpha(\delta)}.$$

Then any sequence $(\delta_n) \to 0$ *has a subsequence* $(\delta_n) \to 0$ *such that* $u_{\delta_n} := u_{\alpha(\delta_n)} \to u_{\mathcal{J}}^\dagger$ *and* $\mathcal{J}(u_\delta) \to \mathcal{J}(u_{\mathcal{J}}^\dagger)$, *where* $u_{\mathcal{J}}^\dagger$ *is a* $\mathcal{J}$*-minimising solution.*

Proof. By lemma 4.30, we know that Φ_α is coercive, and by theorem 4.31, we know there exists a $\mathcal{J}$-minimising solution $u_{\mathcal{J}}^\dagger$. For any (u_{δ_n}) with $\delta_n \to 0$ we find

$$\Phi_\alpha(u_{\delta_n}) \le \Phi_\alpha(u_{\mathcal{J}}^\dagger) = \frac{1}{2}\left\|Au_{\mathcal{J}}^\dagger - f_\delta\right\|^2 + \alpha(\delta)\mathcal{J}(u_{\mathcal{J}}^\dagger) \le \frac{\delta^2}{2} + \alpha(\delta)\mathcal{J}(u_{\mathcal{J}}^\dagger), \tag{6}$$

which is bounded, so by coercivity of Φ_α we conclude that (u_{δ_n}) is bounded.

After taking a subsequence if necessary, we conclude that $u_{\delta_n} \overset{\tau_X}{\to} u_0$. We show that u_0 is a $\mathcal{J}$-minimising solution. Since $A\colon \tau_X \to \tau_Y$ is continuous and $\|\cdot\|_{\mathcal{Y}}$ is $\tau_{\mathcal{Y}}$-LSC, we get

$$\|Au_0 - f\| \le \liminf_{n\to\infty} \|Au_{\delta_n} - f\| \le \liminf_{n\to\infty} \|Au_{\delta_n} - f_{\delta_n}\| + \|f - f_{\delta_n}\| = 0,$$

since $\|Au_{\delta_n} - f\delta_n\| \le 2\Phi_\alpha(u_{\delta_n}) \to 0$ by eq. (6). This shows that u_0 is a least-squares solution.

Furthermore, we have

$$\mathcal{J}(u_0) \le \liminf_{n\to\infty} \mathcal{J}(u_{\delta_n}) \le \limsup_{n\to\infty} \mathcal{J}(u_{\delta_n}) \le \limsup_{n\to\infty} \frac{\Phi_\alpha(u_{\delta_n})}{\alpha(\delta)} \le \limsup_{n\to\infty} \frac{1}{2}\frac{\delta_n^2}{\alpha(\delta_n)} + \mathcal{J}(u_{\mathcal{J}}^\dagger) = \mathcal{J}(u_{\mathcal{J}}^\dagger)$$

which proves that u_0 minimises $\mathcal{J}$.

To prove that u_{δ_n} is actually a minimising sequence, note that all above inequalities are in fact equalities, so $\liminf_{n\to\infty} \mathcal{J}(u_{\delta_n}) = \limsup_{n\to\infty} \mathcal{J}(u_{\delta_n}) = \lim_{n\to\infty} \mathcal{J}(u_{\delta_n}) = \mathcal{J}(u_{\mathcal{J}}^\dagger)$.

□

4.3 Total variation regularisation

Definition 4.33. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded comain and $u \in L^1(\Omega)$. Let

$$\mathcal{D}(\Omega, \mathbb{R}^n) := \left\{\varphi \in C_0^\infty(\Omega, \mathbb{R}^n) \mid \sup_{x\in\Omega} \|\varphi(x)\|_2 \le 1\right\}.$$

We define the *total variation* of $u \in L^1(\Omega)$ as

$$\mathrm{TV}(u) := \sup_{\varphi\in\mathcal{D}(\Omega,\mathbb{R}^n)} \int_\Omega u(x)(\boldsymbol{\nabla}\cdot\varphi)(x)\,\mathrm{d}x\,.$$

Example 4.34. If $u \in L^1(\Omega)$ has a weak derivative $\boldsymbol{\nabla}u \in L^1$ (i.e., $u \in W^{1,1}$), then since any $\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)$ is 0 on the boundary of Ω, we have

$$\mathrm{TV}(u) = \sup_{\varphi\in\mathcal{D}(\Omega,\mathbb{R}^n)} \int_\Omega u(x)(\boldsymbol{\nabla}\cdot\varphi)(x)\,\mathrm{d}x = \sup_{\varphi\in\mathcal{D}(\Omega,\mathbb{R}^n)} \int_\Omega -\boldsymbol{\nabla}u\cdot\varphi\,\mathrm{d}x\,,$$

By Cauchy-Schwarz we have $|\nabla u(x)\cdot\varphi(x)| \le \|\nabla u(x)\|_2\|\varphi(x)\|_2$ which gives an upper bound for $\mathrm{TV}(u)$, and it can be shown that this upper bound is indeed an equality, so

$$\mathrm{TV}(u) = \int_\Omega \|\nabla u\|_2\,\mathrm{d}x = \|\nabla u\|_{L^1}.$$

Proposition 4.35. *The functional* TV *is proper, convex, and absolutely one-homogeneous, and for any constant c and any $u \in L^1(\Omega)$ we have*

$$\mathrm{TV}(c) = 0, \quad \mathrm{TV}(u + c) = \mathrm{TV}(u).$$

Proof. For any test function φ we have

$$\int_\Omega \boldsymbol{\nabla} \cdot \varphi \, \mathrm{d}x = \int_{\partial\Omega} \varphi \cdot \mathbf{n} \, \mathrm{d}x = 0$$

since $\varphi = 0$ on the boundary. This shows immediately that $\mathrm{TV}(c) = 0$ and that $\mathrm{TV}(u+c) = \mathrm{TV}(u)$. □

Remark. In fact, the converse also holds: $\mathrm{TV}(u) = 0$ if and only if u is constant.

Example 4.36. Let $C \subseteq \Omega$ be a bounded domain with smooth boundary. Then we have

$$\mathrm{TV}(\mathbb{1}_C) = \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_C \boldsymbol{\nabla} \cdot \varphi \, \mathrm{d}x = \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_{\partial C} \varphi \cdot \mathbf{n} \, \mathrm{d}s \leq \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_{\partial C} \|\varphi\|_2 \|\mathbf{n}\|_2 \, \mathrm{d}s = \int_{\partial C} 1 \, \mathrm{d}s \,,$$

and we can show that the inequality here is in fact an equality, i.e., $\mathrm{TV}(\mathbb{1}_C) = \int_{\partial C} 1 \, \mathrm{d}s$, which is the perimeter of C.

Definition 4.37. We define the space of functions of *bounded variation* as

$$\mathrm{BV}(\Omega) \coloneqq \big\{ u \in L^1(\Omega) : \|u\|_{\mathrm{BV}} \coloneqq \|u\|_{L^1} + \mathrm{TV}(u) < \infty \big\}.$$

Remark. It can be shown that BV is the dual of a separable Banach space, and that weak-$*$ convergence in BV is equivalent to strong convergence in L^1 and convergence of the values $\mathrm{TV}(u_n) \to \mathrm{TV}(u)$.

Definition 4.38. We define a *domain* in $\mathbb{R}^n$ as an open connected subset, and it is called a *Lipschitz domain* if its boundary is locally the graph of a Lipschitz continuous function (see Wikipedia for the details).

Theorem 4.39. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded Lipschitz domain and $p, m \in \mathbb{N}$. Define*

$$p^* \coloneqq \begin{cases} \frac{np}{n-mp}, & n > mp, \\ \infty, & n \leq mp. \end{cases}$$

Then the embedding $W^{m,p} \to L^q(\Omega)$ is continuous for all $q \in [1, p^*]$ and compact for all $q \in [1, p^*)$.

A consequence ($p = m = 1$) is the following:

Corollary 4.40. *For any bounded Lipschitz domain $\Omega \subseteq \mathbb{R}^n$, the* $\mathrm{BV}(\Omega)$ *is compactly embedded in $L^1(\Omega)$ for $n \geq 2$. For $n = 2$, the embedding* $\mathrm{BV}(\Omega) \to L^2(\Omega)$ *is continuous.*

Theorem 4.41. *Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded. Then* TV *is LSC in $L^1(\Omega)$.*

Proof. Let $(u_j) \subseteq \mathrm{BV}(\Omega)$ converge to $u \in L^1(\Omega)$. Then for any test function we have

$$\begin{aligned} \mathrm{TV}(u) &= \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_\Omega u_j(x) (\boldsymbol{\nabla} \cdot \varphi)(x) \, \mathrm{d}x \\ &= \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \lim_{j \to \infty} \int_\Omega u_j(x) (\boldsymbol{\nabla} \cdot \varphi)(x) \, \mathrm{d}x \\ &\leq \liminf_{j \to \infty} \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_\Omega u_j(x) (\boldsymbol{\nabla} \cdot \varphi)(x) \, \mathrm{d}x \\ &= \liminf_{j \to \infty} \mathrm{TV}(u_j). \end{aligned}$$

□

TODO: what is the point/use of proposition 4.3.12 and corollary 4.3.13?

Theorem 4.42. *Let $\mathcal{X} = L^1(\Omega)$, where $\Omega \subseteq \mathbb{R}^n$ is bounded Lipschitz, and $\mathcal{Y}$ a Banach space. Let $A \in \mathcal{B}(L^1, \mathcal{Y})$ such that $A(1) \neq 0$. Then u_δ converges strongly in L^1 (along a subsequence) to a TV-minimising solution as $\delta \to 0$ if $\alpha(\delta)$ is chosen as required by theorem 4.32.*

Proof. We need to check that the conditions of theorem 4.32 are satisfied. We have assumed that $\mathcal{R}(A)$ is dense, so $\inf_{v \in \mathcal{X}} \|Av - f\| = 0$. Since $\mathrm{BV}(\Omega)$ is compactly embedded in $L^1(\Omega)$, we know that sequences with a finite TV-value have (strongly) convergent subsequences. We have checked that TV is convex and LSC, that $\dim(\mathcal{N}(\mathrm{TV})) < \infty$, and by assumption we have $\mathcal{N}(A) \cap \mathcal{N}(\mathrm{TV}) = \{0\}$. Finally, TV is absolutely one-homogeneous, so all assumptions are valid. □

5 Convex Duality

We are looking for a convergence rate

$$D_{\mathcal{J}}^{\text{symm}}(u_\delta, u_{\mathcal{J}}^\dagger) \le \psi(\delta) \qquad \psi\colon \mathbb{R}_+ \to \mathbb{R}_+, \quad \lim_{\delta\to 0}\psi(\delta) = 0.$$

We will assume for simplicity that $\mathcal{J}$ is absolutely one-homogeneous. We define

$$E(y) \coloneqq \frac{1}{2}\|y - f\|_{\mathcal{Y}}^2, \quad F(u) \coloneqq \alpha\mathcal{J}(u),$$

so that $u_\delta \in \arg\min_{u\in\mathcal{X}} E(u) + F(Au)$.

Lemma 5.1. *Let X be a Banach space and $\varphi(x) = \frac{1}{2}\|x\|_X^2$. Then the convex conjugate of φ is*

$$\varphi^*(\xi) = \frac{1}{2}\|\xi\|_{X^*}^2.$$

Proof. We have

$$\varphi^*(\xi) = \sup_{x\in X}\langle\xi, x\rangle - \frac{1}{2}\|x\|^2 \le \sup_{x\in X}\|x\|\|\xi\| - \frac{1}{2}\|x\|^2.$$

The right-hand is a quadratic function in $\|x\|$ with maximum $\frac{1}{2}\|\xi\|^2$. The inequality is in fact equality (TODO: show) which proves the result. □

Now we compute

$$E^*(\eta) = \langle\eta, f\rangle + \frac{1}{2}\|\eta\|_{\mathcal{Y}^*}, \quad F^*(p) = \chi_{\partial\mathcal{J}(0)}(p/\alpha),$$

which gives the dual problem

$$\sup_{\eta\in\mathcal{Y}^*} -\langle\eta, f\rangle - \frac{1}{2}\|\eta\|^2 - \chi_{\partial\mathcal{J}(0)}\left(-\frac{A^*\eta}{\alpha}\right),$$

which is equivalent to

$$\sup_{\mu\in\mathcal{Y}^*, A^*\mu\in\partial\mathcal{J}(0)} \alpha\left(\langle\mu, f\rangle - \frac{\alpha}{2}\|\mu\|^2\right)$$

Plugging this into theorem 4.25 gives optimality conditions

$$A^*\mu_\delta \in \partial\mathcal{J}(u_\delta), \quad -\alpha\mu_\delta \in \partial\left(\frac{1}{2}\|\cdot\|_{\mathcal{Y}}^2\right)(Au_\delta - f_\delta)$$

TODO: Finish this section (and actually understand what the **** is going on)

6 Bayesian probability and statistics

6.1 Motivation

Let $\mathcal{X}, \mathcal{Y}$ be separable Banach spaces and $A\colon \mathcal{X} \to \mathcal{Y}$ a measurable operator (possibly nonlinear). We consider the inverse problem

$$\text{Find } u \in \mathcal{X} \text{ such that } A(u) + n = f_n,$$

where $f_n \in \mathcal{Y}$ is data and $n \in \mathcal{Y}$ is noise.

In Bayesian probability theory, we attempt to give an estimate for u while also commenting on the uncertainty left in u after measurement. The main idea is to **model uncertain parameters as random variables**. In this case, we assume the noise n and the parameter u are random variables N, U. The distribution of N is usually known, while the distribution of U (called the *prior distribution* $\mu_0 = \mathbb{P}(U \in \cdot)$) describes our prior knowledge of u. We then observe data f_n, which is described as an occurrence of the event $\{f_n = A(U) + N\}$. After our observation, we replace the distribution μ_0 by the distribution

$$\mu_{\text{post}} \coloneqq \mathbb{P}(U \in \cdot \mid f_n = A(U) + N).$$

6.2 Measure theory

Definition 6.1. Let Ω be a topological space, and let $\mathcal{O}$ be the topology on Ω. Then the σ-algebra generated by $\mathcal{O}$ is denoted $\mathcal{B}\Omega$ and called the *Borel σ-algebra* on Ω.

Note that we do not explicitly state the topology used when writing $\mathcal{B}\Omega$, but this is always clear from context.

Definition 6.2. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $(\Omega', \mathcal{F}')$ a measurable space, and $g\colon (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ measurable. Then, the map $\mu \circ g^{-1} = \mu(g \in \cdot)$ defines a measure on $(\Omega', \mathcal{F}')$ called the *pushforward measure*.

Note that if μ is a probability measure, then so is $\mu(g \in \cdot)$.

Definition 6.3. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $g\colon (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ be measurable and non-negative. Then the map

$$\nu\colon \mathcal{F} \to [0, \infty]\colon F \mapsto \int_F g \,\mathrm{d}\mu$$

defines a measure on $(\Omega, \mathcal{F})$, called the *measure with (μ-)density g*. If ν is a probability measure, g is called a *(μ-)probability density*.

Definition 6.4. Let $(\Omega, \mathcal{F})$ be a measurable space and μ, ν two measures on Ω. Then ν is called *absolutely continuous* with respect to μ, denoted $\nu \ll \mu$, if

$$\mu(F) = 0 \implies \nu(F) = 0 \qquad \text{for all } F \in \mathcal{F}.$$

Theorem 6.5 (Radon-Nikodym)**.** *Let $(\Omega, \mathcal{F})$ be a measurable space and let μ, ν be σ-finite measures on $(\Omega, \mathcal{F})$. Then, the following are equivalent:*

1. *$\nu \ll \mu$;*
2. *There is a measurable $g\colon (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$, with*

$$\nu(F) = \int_F g \,\mathrm{d}\mu \qquad \textit{for all } F \in \mathcal{F}.$$

The function g is μ-a.e. unique, and is called the Radon-Nikodym derivative, denoted $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

6.3 Conditional probability

We will consider $(\Omega, \mathcal{F}, \mathbb{P})$ as underlying probability space for any random variable, and assume Ω is separable and completely metrisable (this is also called *Polish*) and $\mathcal{F} = \mathcal{B}\Omega$. We also write

$$\mathbb{E}[\varphi] \coloneqq \int_\Omega \varphi \, \mathrm{d}\mathbb{P}$$

for any $\varphi \in L^1(\Omega)$.

Theorem 6.6. *Let $U\colon (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ and $Y\colon (\Omega, \mathcal{F}) \to (\mathcal{Y}, \mathcal{B}\mathcal{Y})$ be random variables and let U be integrable. Then there exists a measurable function $h\colon (\mathcal{Y}, \mathcal{B}\mathcal{Y}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ such that*

$$\int_F h(y)\mathbb{P}(Y \in \mathrm{d}y) = \int_{\{Y \in F\}} U \, \mathrm{d}\mathbb{P} \qquad \text{for all } F \in \mathcal{F}.$$

The function h is $\mathbb{P}(Y \in \cdot)$-a.s. unique.

Definition 6.7. For $y \in \mathcal{Y}$, we call $h(y)$ (from the previous theorem) the conditional expectation of U given $Y = y$, denoted $h(y) =: \mathbb{E}[U \mid Y = y]$ (which is well-defined up to $\mathbb{P}(Y \in \cdot)$ null sets).

Definition 6.8. Let $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ be measurable spaces and $M\colon \Omega \times \mathcal{F}' \to [0, 1]$. Then M is called a *Markov kernel* if

1. $M(\omega, \cdot)$ is a probability measure for all $\omega \in \Omega$;
2. $M(\cdot, F')$ is measurable for all $F' \in \mathcal{F}'$.

It can be shown that $(y, F) \mapsto \mathbb{E}[\mathbb{1}_F \mid Y = y]$ (which represents the probability of F given $Y = y$) is a Markov kernel. This is clarified in the following theorem:

Theorem 6.9. *Let $U\colon (\Omega, \mathcal{F}) \to (\mathcal{X}, \mathcal{B}\mathcal{X})$ and $Y\colon (\Omega, \mathcal{F}) \to (\mathcal{Y}, \mathcal{B}\mathcal{Y})$ be random variables. Then there exists a Markov kernel from $(\mathcal{Y}, \mathcal{B}\mathcal{Y})$ to $(\mathcal{X}, \mathcal{B}\mathcal{X})$ with*

$$\int_F M(y, F')\mathbb{P}(Y \in \mathrm{d}y) = \mathbb{P}(\{Y \in F\} \cap \{U \in F'\}) \qquad \text{for all } F \in \mathcal{B}\mathcal{Y} \text{ and } F' \in \mathcal{B}\mathcal{X}.$$

Furthermore, M is $\mathbb{P}(Y \in \cdot)$-a.s. unique, and it is called the conditional probability distribution of U given $Y = y$. We write $M(y, F) \coloneqq \mathbb{P}(U \in F \mid Y = y)$.

Furthermore, any Markov kernel represents the conditional probability measure of some random variable:

Theorem 6.10. *Let $M\colon \Omega' \times \mathcal{F}'' \to [0, 1]$ be a Markov kernel from $(\Omega', \mathcal{F}')$ to $(\Omega'', \mathcal{F}'')$. Then there is an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variables $X'\colon \Omega \to \Omega'$, $X''\colon \Omega \to \Omega''$ such that*

$$M(\omega', F'') = \mathbb{P}(X'' \in F'' \mid X' = \omega') \qquad \text{for all } F'' \in \mathcal{F}'' \text{ and } \mathbb{P}(X' \in \cdot)\text{-almost all } \omega' \in \Omega'.$$

We now consider the marginal density function of jointly distributed random variables:

Lemma 6.11. *Let U, Y be random variables with joint distribution $\mathbb{P}((U, Y) \in \cdot)$, which is absolutely continuous w.r.t. a σ-finite measure ν on $(\mathcal{X} \times \mathcal{Y}, \mathcal{B}\mathcal{X} \otimes \mathcal{B}\mathcal{Y})$. Assume that $\nu = \nu_U \otimes \nu_Y$ and that $(\mathcal{X}, \mathcal{B}\mathcal{X}, \nu_U)$ and $(\mathcal{Y}, \mathcal{B}\mathcal{Y}, \nu_Y)$ are σ-finite. Writing $g_{U,Y} \coloneqq \frac{\mathrm{d}\mathbb{P}((U,Y) \in \cdot)}{\mathrm{d}\nu}$, we have*

$$\mathbb{P}(U \in \cdot) \ll \nu_U, \quad \mathbb{P}(Y \in \cdot) \ll \nu_Y,$$

with probability density functions

$$g_U \coloneqq \int_{\mathcal{Y}} g_{U,Y} \, \mathrm{d}\nu_Y = \frac{\mathrm{d}\mathbb{P}(U \in \cdot)}{\mathrm{d}\nu_U},$$
$$g_Y \coloneqq \int_{\mathcal{X}} g_{U,Y} \, \mathrm{d}\nu_U = \frac{\mathrm{d}\mathbb{P}(Y \in \cdot)}{\mathrm{d}\nu_Y}.$$

We can use the above theorem to consider the density function of U conditional on Y:

Theorem 6.12. *Under the assumptions of the previous lemma, we have $\mathbb{P}(U \in \cdot \mid Y = y) \ll \nu_U$, and its ν_U density is*

$$g_{U|Y=y}(u) = \mathbb{1}_{g_Y(y)>0} \frac{g_{U,Y}(u,y)}{g_Y(y)},$$

which is uniquely defined up to null sets.

Definition 6.13. Let $g_U, g_Y, g_{U,Y}, g_{U|Y=y}, g_{Y|U=u}$ be the probability densities from the previous theorem. Then g_U is called the *marginal probability density* of U, $g_{U,Y}$ is called the *joint probability density* of U and Y, and $g_{U|Y=y}$ is called the *conditional density* of U given $Y = y$.

6.4 Bayesian statistics

6.4.1 Statistical models

Definition 6.14. Let $\mathcal{X}, \mathcal{Y}$ be separable Banach spaces. We call $\mathcal{X}$ the *parameter space* and $\mathcal{Y}$ the *data space*. Let $\mathcal{P} \coloneqq \{M(\cdot \mid u) : u \in \mathcal{X}\}$, where M is a Markov kernel from $(\mathcal{X}, \mathcal{B}\mathcal{X})$ to $(\mathcal{Y}, \mathcal{B}\mathcal{Y})$. The tuple $(\mathcal{Y}, \mathcal{P})$ is called the *statistical model*.

Usually, we distinguish between *parametric models* (where $\mathcal{X}$ is finite-dimensional) and *nonparametric models*. However, in Bayesian statistics, this distinction rarely matters.

Given a "true" parameter $u^* \in \mathcal{X}$, a distribution $Y \sim M(\cdot \mid u^*)$ and a realisation y of Y, we aim to find u^* based on y. The probability measure $M(\cdot \mid u^*)$ is called the *data-generating distribution*.

Definition 6.15. Let $(\mathcal{Y}, \mathcal{P})$ be a statistical model and $L \colon (\mathcal{X} \times \mathcal{Y}, \mathcal{B}\mathcal{X} \otimes \mathcal{B}\mathcal{Y}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ such that

$$\mathcal{P} = \left\{ F \mapsto \int_F L(y \mid u) \mathrm{d}\mu(y) : u \in \mathcal{X} \right\}$$

for some measure μ on $(\mathcal{Y}, \mathcal{B}\mathcal{Y})$. Then L is called the *(data) likelihood*.

The likelihood is a conditional density $g_{Y|U=u}$ for some random variable U.

6.4.2 Bayes' formula

Assume we are given a likelihood $L = f_{Y|U=u}$ for some U. We are interested in finding $f_{U|Y=y}$, the posterior density function of U. For this we use Bayes's formula:

Theorem 6.16 (Bayes)**.** *Let U, Y be random variables as in theorem 6.12. Then we have*

$$g_{U|Y=y}(u) = \frac{g_{Y|U=u}(y) g_U(u)}{g_Y(y)}$$

for $u \in \mathcal{X}$, ν_U-a.e., and $y \in \mathcal{Y}$, $\mathbb{P}(Y \in \cdot)$-a.e. with $g_Y(y) > 0$.

Definition 6.17. In the formulation of Bayes' formula:

1. $Z(y) \coloneqq g_Y(y)$ is called the *model evidence* or *marginal likelihood*;
2. $L(y \mid u) \coloneqq g_{Y|U=u}(y)$ is called the *(data) likelihood*;
3. $\mu_0 \coloneqq \mathbb{P}(U \in \cdot)$ is called the *prior measure*;
4. $\mu_{\text{post}} \coloneqq \mathbb{P}(U \in \cdot \mid Y = y)$ is called the *posterior measure*.

If we assume that $\nu_U \coloneqq \mu_0$ (we can choose ν_U freely so long as μ_0 has a probability density w.r.t. ν_0), we obtain from Bayes' formula the formulation

$$\frac{\mathrm{d}\mu_{\mathrm{post}}}{\mathrm{d}\mu_0} = \frac{L(y \mid u)}{Z(y)} \qquad (\mu_0\text{-a.s.}).$$

If μ_0 has a density with respect to some other measure ν (say, Lebesgue measure), then we obtain

$$\frac{\mathrm{d}\mu_{\mathrm{post}}}{\mathrm{d}\nu} = \frac{\mathrm{d}\mu_{\mathrm{post}}}{\mathrm{d}\mu_0}\frac{\mathrm{d}\mu_0}{\mathrm{d}\nu} = \frac{L(y \mid u)}{Z(y)}\frac{\mathrm{d}\mu_0}{\mathrm{d}\nu}.$$

In practice, this means that computing the density of the posterior measure comes down to multiplying the likelihood with the prior measure and normalising accordingly.

7 Bayesian inverse problems

7.1 Bayesian inverse problems

Definition 7.1. Let $(\Omega', \mathcal{F}')$ be some measurable space. We denote the space of probability measures on $(\Omega', \mathcal{F}')$ by $\mathrm{Prob}(\Omega', \mathcal{F}')$. Furthermore, if ν is a σ-finite measure on $(\Omega', \mathcal{F}')$, we let $\mathrm{Prob}(\Omega', \mathcal{F}', \nu)$ denote the space of probability measures $\mu \in \mathrm{Prob}(\Omega', \mathcal{F}')$ with $\mu \ll \nu$.

Definition 7.2. Let $\mu_0 \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ and $L\colon (\mathcal{X}\times\mathcal{Y}, \mathcal{B}\mathcal{X}\otimes\mathcal{B}\mathcal{Y}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ be a measurable function. Then we define the *Bayesian inverse problem with prior μ_0 and likelihood L* to be the problem of finding $\mu_{\mathrm{post}} \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ with

$$\frac{\mathrm{d}\mu_{\mathrm{post}}}{\mathrm{d}\mu_0}(u) = \frac{L(f_n \mid u)}{\int_{\mathcal{X}} L(f_n \mid u)\,\mathrm{d}\mu_0(u)} \qquad (u \in \mathcal{X};\, \mu_0\text{-a.s.}),$$

for any data set $f_n \in \mathcal{Y}$.

Definition 7.3. Consider the BIP as in the previous definition, let $P \subseteq \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ be a space of probability measures and d a metric on P. Then the problem is (P, d)*-well-posed* if

1. For all $f_n \in \mathcal{Y}$, there exists a solution $\mu_{\mathrm{post}} \in P$ (existence);
2. For all $f_n \in \mathcal{Y}$, the solution $\mu_{\mathrm{post}} \in P$ is unique (uniqueness);
3. The map $f_n \mapsto \mu_{\mathrm{post}}$ is continuous from $(\mathcal{Y}, \mathcal{B}\mathcal{Y})$ to (P, d) (stability).

If $\int_{\mathcal{X}} L(f_n \mid u)\mathrm{d}\mu_0(u) \in (0, \infty)$, then existence and uniqueness is clear. This is immediate if $L > 0$ and $L(f_n \mid \cdot) \in L^1$ for all f_n.

7.2 Metrics on spaces of probability measures

We consider two metrics on spaces of probability measures:

Definition 7.4. Let $(\Omega', \mathcal{F}')$ be a measurable space.

1. We define the *total variation distance* on $\mathrm{Prob}(\Omega', \mathcal{F}')$ by
$$d_{\mathrm{TV}}(\mu, \nu) \coloneqq \sup_{F' \in \mathcal{F}'} |\mu(F') - \nu(F')|.$$
2. Suppose that Ω' is a topological space and $\mathcal{F}' = \mathcal{B}\Omega'$. Let $(\mu_n) \subseteq \mathrm{Prob}(\Omega', \mathcal{F}')$ and $\mu \in \mathrm{Prob}(\Omega', \mathcal{F}')$. We say that μ_n converges *weakly* to μ as $n \to \infty$, denoted $\mu_n \rightharpoonup \mu$, if
$$\int_{\Omega'} g\,\mathrm{d}\mu_n \to \int_{\Omega'} g\,\mathrm{d}\mu \quad \text{for all continuous bounded functions } g\colon (\Omega', \mathcal{B}\Omega') \to (\mathbb{R}, \mathcal{B}\mathbb{R}).$$

It is known that the weak convergence topology is metrisable, however, the metric itself is hardly used. We will denote this metric d_{LP}.

Lemma 7.5. *Convergence in TV-distance is stronger than weak convergence.*

Lemma 7.6. *Let $\mu, \nu \in \mathrm{Prob}(\Omega, \mathcal{F})$, and let ρ be a σ-finite measure with $\mu, \nu \ll \rho$. Then we have*

$$d_{\mathrm{TV}}(\mu, \nu) = \frac{1}{2}\int_{\Omega} \left|\frac{\mathrm{d}\mu}{\mathrm{d}\rho} - \frac{\mathrm{d}\nu}{\mathrm{d}\rho}\right| \mathrm{d}\rho\,.$$

One can always choose $\rho = \mu + \nu$ in the previous lemma.

7.3 Stability

We will assume the following, given a Bayesian inverse problem with prior μ_0 and likelihood L:

A1 $L(\cdot \mid u)$ is strictly positive;

A2 $L(f_n \mid \cdot) \in L^1(\mu_0)$;

A3 There exists $h \in L^1(\mu_0)$ such that $L(f_n' \mid \cdot) \leq h(\cdot)$ for all $f_n' \in \mathcal{Y}$;

A4 $L(\cdot \mid u)$ is continuous.

Note that A1 and A2 were already required for existence and uniqueness. A4 asks for continuity only in the data, not in the parameter space.

Theorem 7.7. *Given a BIP with prior μ_0 and likelihood L which satisfies assumptions (A1)-(A4) and let $P = \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$. Then the problem is (P, d_{TV})-well-posed, and therefore also (P, D_{LP})-well-posed.*

Proof. We have already proved existence and uniqueness. By lemma 7.6, to prove that $f_n \mapsto \mu_{\mathrm{post}}$ is continuous, it suffices to prove for any $f_n^{(m)} \to f_n$ that

$$\int_{\mathcal{X}} \left| \frac{L(f_n \mid u)}{Z(f_n)} - \frac{L(f_n^{(m)} \mid u)}{Z(f_n^{(m)})} \right| \mathrm{d}\mu_0(u) \to 0.$$

For this, we first show that the map $f_n \mapsto L(f_n \mid \cdot)$ is continuous as a map from $\mathcal{Y}$ to $L^1(\mu_0)$. Note that

$$\lim_{m \to \infty} \int_{\mathcal{X}} \left| L(f_n^{(m)} \mid u) - L(f_n \mid u) \right| \mathrm{d}\mu_0(u) = 0$$

since we can exchange limit and integral by the dominated convergence theorem (since $L \in L^1$, the integrand is bounded by $2|L|$), and then apply assumption A4 which guarantees continuity of $L(\cdot \mid u)$.

Analogously, we can show that the map $f_n \mapsto Z(f_n)$ is continuous, and therefore, the map $f_n \mapsto L(f_n \mid \cdot)/Z(f_n)$ is continuous, from which the claim follows. □

8 Function space priors and Monte Carlo

This section covers two different topics:

1. The general definition of a Gaussian distribution on an infinite-dimensional space (useful for constructing priors);
2. Description of Monte Carlo methods for approximating the posterior.

8.1 Gaussian measures

Recall that a random variable in $\mathbb{R}^d$ is called Gaussian if and only if every linear combination of its components is a Gaussian random variable on $\mathbb{R}$. This can be generalised to any separable Banach space $(\mathcal{X}, \mathcal{B}\mathcal{X})$:

Definition 8.1. Let μ be a probability measure on $\mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ and let $U \sim \mu$. We call μ *Gaussian* if, for all $\ell \in \mathcal{X}^*$, the random variable $\langle \ell, U \rangle$ has a normal distribution.

We define the *mean* of μ by

$$a_\mu \in \mathcal{X}^{**} \colon \ell \mapsto \int_{\mathcal{X}} \langle \ell, u \rangle \, \mathrm{d}\mu \, (u),$$

and the *covariance operator* of μ by $R_\mu \colon (\mathcal{X}^*)^2 \to \mathbb{R}$, where

$$R_\mu(\ell, \ell') = \int_{\mathcal{X}} (\langle \ell, u \rangle - a_\mu(\ell))(\langle \ell', u \rangle - a_\mu(\ell')) \, \mathrm{d}\mu(u)$$

If $\mathcal{X}$ is a function space, U is called a *Gaussian random field*.

This definition is not constructive, and in general, it is nontrivial to construct a Gaussian random field. On a separable Hilbert space however, an explicit construction is possible.

Definition 8.2. Let $\mathcal{X}$ be a separable Hilbert space and $C \in \mathcal{K}(\mathcal{X}, \mathcal{X})$ self-adjoint, and write the eigendecomposition

$$Cx = \sum_{i=1}^{\infty} \lambda_i \langle x, \varphi_i \rangle \varphi_i$$

where $|\lambda_1| \geq |\lambda_2| \geq \cdots$. Then C is called a *trace class* operator if $(\lambda_i)_{i \in \mathbb{N}} \in \ell^1$.

Proposition 8.3. *Let $\mathcal{X}$ be a separable Hilbert space and C a positive semi-definite trace class operator with an eigenvalue decomposition as in the previous definition. Letting $m \in \mathcal{X}$ and $\xi \sim \mathrm{N}(0,1)^{\otimes \mathbb{N}}$ (a sequence of independent $\mathrm{N}(0,1)$ random variables), we have that*

$$U \coloneqq m + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i \varphi_i.$$

is distributed according to a Gaussian measure with mean m and covariance operator C (in the sence that $R_\mu(x, y) = \langle Cx, y \rangle$).

Proof. Let $k \in \mathbb{N}$ and $U_k \coloneqq m + \sum_{i=1}^{k} \sqrt{\lambda_i} \xi_i \varphi_i$, and let $x \in \mathcal{X}$ and $x_i \coloneqq \langle x, \varphi_i \rangle$. Then we have

$$\begin{aligned} \langle x, U_k \rangle &= \langle x, m \rangle + \sum_{i=1}^{k} \sqrt{\lambda_i} \langle x, \varphi_i \rangle \xi_i \\ &= \langle x, m \rangle + \sum_{i=1}^{k} \sqrt{\lambda_i} x_i \xi_i, \end{aligned}$$

and since $\sqrt{\lambda_i} x_i \xi_i \sim \mathrm{N}(0, \lambda_i x_i^2)$, it can be shown that $\langle x, U_k \rangle$ converges weakly to the distribution $\mathrm{N}(\langle x, m \rangle, \sum_{i=1}^{\infty} \lambda_i x_i^2)$ if the sum $\sum_{i=1}^{\infty} \lambda_i x_i^2$ is finite.

Indeed, the sum is finite: $(\lambda_i) \in \ell^1$ by assumption and $(x_i^2) \in \ell^1$ since $\sum_i x_i^2 = \|x\|^2 < \infty$. It is known that the pointwise product of two ℓ^1 sequences lies again in ℓ^1.

This proves that U has a normal distribution, and we cam compute the mean and covariance using Fubini's theorem, which proves the claim. □

Example 8.4. Let $D = [0,1]^2$ and $\mathcal{X} \coloneqq L^2(D, \mathcal{B}D, \lambda_2)$ and $\ell > 0$, $\sigma^2 \geq 0$. Then the *exponential covariance function* is defined as

$$c_{\exp}(x, y) \coloneqq \sigma^2 \exp(-\|x - y\|_2 / \ell),$$

while the *Gaussian covariance function* is defined as

$$c_{\mathrm{N}}(x, y) \coloneqq \sigma^2 \exp\bigl(-\|x - y\|_2 / (2\ell^2)\bigr).$$

The parameter ℓ is called *correlation length* while σ^2 is called *pointwise variance*.

8.2 Monte Carlo methods

Generally, Monte Carlo techniques aim at approximating the integral $\bar{g} \coloneqq \int_{\mathcal{X}} g \, \mathrm{d}\mu$ by generating independent samples $U_1, U_2, \ldots \overset{\text{iid}}{\sim} \mu$ and computing the estimator

$$\bar{g} \approx \hat{g}_M \coloneqq \frac{1}{M} \sum_{m=1}^{M} g(U_m).$$

By the strong law of large numbers, we know that $\hat{g}_m \to \bar{g}$ for $M \to \infty$, $\mathbb{P}$-almost surely. Furthermore, if $\mathrm{Var}_\mu(g) < \infty$, then we have the following expression for the standard deviation:

$$\sqrt{\mathbb{E}[(\hat{g}_m - \bar{g})^2]} = \sqrt{\mathrm{Var}_\mu(g)} \cdot M^{-1/2},$$

which gives a convergence rate of $M^{-1/2}$.

This rate is not the best possible rate: if we assume that g is smooth in some sense, much better rates can be obtained. However, a large advantage is that the rate $M^{-1/2}$ is independent of the dimension of the domain.

The standard Monte Carlo method does not work for Bayesian inverse problems because in general it is not possible to sample independently from the posterior measure. We must therefore sample dependently (there are also other options such as sampling independently from a different measure and correct using weights).

Definition 8.5. Let $(U_n)_{n=1}^{\infty}$ be a sequence of $\mathcal{X}$-valued random variables. Then (U_n) is called a *Markov chain* if, for any $n \in \mathbb{N}$ and $u_1, \ldots, u_n \in \mathcal{X}$, we have

$$\mathbb{P}(U_{n+1} \in \cdot \mid U_1 = u_1, \ldots, U_n = u_n) = \mathbb{P}(U_{n+1} \in \cdot \mid U_n = u_n)$$

We call the chain *time-homogeneous* if, for all $u \in \mathcal{X}, n \in \mathbb{N}$ we have

$$\mathbb{P}(U_{k+2} \in \cdot \mid U_{k+1} = u) = \mathbb{P}(U_2 \in \cdot \mid U_1 = u).$$

A time-homogeneous Markov chain can be fully represented by a Markov kernel $K \colon \mathcal{B}\mathcal{X} \times \mathcal{X} \to [0,1]$ where

$$K(B \mid u) \coloneqq \mathbb{P}(U_2 \in B \mid U_1 = u)$$

Let $\mu \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ be any probability meaure, then we denote the *composition* of μ and K by

$$\mu K(B) \coloneqq \int_{\mathcal{X}} K(B \mid u) \, \mathrm{d}\mu(u) \qquad (B \in \mathcal{B}\mathcal{X}),$$

and we call μ *stationary* w.r.t. K if $\mu K = \mu$. Note that μ is stationary w.r.t. K if and only if $U_1 \sim \mu \implies U_2 \sim \mu$.

Finally, we say that the Markov kernel K satisfies *detailed balance* w.r.t. $\mu' \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ if

$$\int_B K(A \mid u) \, \mathrm{d}\mu'(u) = \int_A K(B \mid u) \, \mathrm{d}\mu'(u) \qquad (A, B \in \mathcal{B}\mathcal{X}).$$

Lemma 8.6. *Let $K \colon \mathcal{B}\mathcal{X} \times \mathcal{X} \to [0,1]$ be a Markov kernel that satisfies detailed balance w.r.t. $\mu \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$. Then K is stationary w.r.t. μ.*

Definition 8.7. Let $\mu \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X}, \nu)$ and let $g \colon (\mathcal{X}, \mathcal{B}\mathcal{X}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ be positive with $g = c\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$ for some $c > 0$. Moreover, let $Q \colon \mathcal{X} \times \mathcal{B}\mathcal{X} \to [0,1]$ be a Markov kernel, given by a positive function $q \colon (\mathcal{X}, \mathcal{B}\mathcal{X})^2 \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ with

$$Q(A \mid u) = \int_A q(u' \mid u) \, \mathrm{d}\nu(u') \qquad (A \in \mathcal{B}\mathcal{X}, u \in \mathcal{X}).$$

Now the *Metropolis-Hastings* Markov kernel is given by

$$K_{\mathrm{MH}}(A \mid u) \coloneqq \delta(A - u) \int_{\mathcal{X}} (1 - \alpha(u, u''))Q(\mathrm{d}u'' \mid u) + \int_A \alpha(u, u')Q(\mathrm{d}u' \mid u) \qquad (u \in \mathcal{X}, A \in \mathcal{B}\mathcal{X}),$$

where $\alpha(u, u') \coloneqq 1 \vee \frac{g(u')q(u|u')}{g(u)q(u'|u)}$.

The *Metropolis-Hastings MCMC method* (MCMC = Markov Chain Monte Carlo) is given as follows:

1. Start with some initial $U_1 \in \mathcal{X}$ and set $m = 1$;
2. Sample $U^* \sim Q(\cdot \mid U_m)$;
3. With probability $\alpha(U_m, U^*)$ set $U_{m+1} \leftarrow U^*$, otherwise set $U_{m+1} \leftarrow U_m$.
4. Increment $m \leftarrow m + 1$ and go to step 2.