1. Metric spaces and Banach's FPT (Fixed Point Theorem)

Def 1.1 (Metric and Metric Space)

Let X be a nonempty set, and $d : X^2 \to \mathbb{R}$ be a function satisfying:

- $d(x, y) = 0$ iff $x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) \leq d(x, z) + d(z, y)$
  Then the function d is called the metric, the pair (X, d) is called the metric space, and the number $d(x, y)$ is called the distance between x and y in X.

Examples

- $(\mathbb{R}^n, d)$, with $d(x, y) = \left(\sum_{i=1}^{n}(x_i - y_i)^2\right)^{\frac{1}{2}}$
- $(\mathbb{R}^n, d)$, with $d(x, y) = \max\limits_{i \leq n} |x_i - y_i|$
- $(C[a, b], d)$, with $d(f, g) = \left(\int_a^b |f(x) - g(x)|^2 dx\right)^{\frac{1}{2}}$
- $(C[a, b], d)$, with $d(f, g) = \sup\limits_{a \leq x \leq b} |f(x) - g(x)|$
- $(L^p[a, b], d)$, with $d(f, g) = \left(\int_a^b |f(x) - g(x)|^p dx\right)^{\frac{1}{p}}$

Def 1.2 (Fixed Point)

A fixed point of the mapping $T : X \to X$ is the point $x^* \in X$ such that $T(x^*) = x^*$.

Def 1.3 (Contraction)

Let (X, d) be a metric space. A mapping $T : X \to X$ is called a contraction on X of there exists a constant $0 < k < 1$ such that

$$d(T(x), T(y)) \leq kd(x, y)$$

for all $x, y \in X$.

Theorem 1.1 (Banach's FPT)

Let (X, d) be a complete metric space and let $T : X \to X$ be a contraction on X. Then T has a unique fixed point $x^* \in X$.

Corollary 1.1 (Banach's FPT)

The iterative sequence $x_{n+1} = T(x_n)$ for $n = 1, 2, \ldots$ with arbitrary starting point $x_0 \in X$ converges, under assumptions of Banach's FPT, to the unique fixed point of T. Moreover, the following estimates hold:

- $d(x_m, x^*) \leq \frac{k^m}{1-k} d(x_1, x_0)$ - the prior estimate,
- $d(x_m, x^*) \leq \frac{k}{1-k} d(x_{m-1}, x_m)$ - the posterior estimate.

2. Applications of Banach's FPT

2.1 Applications to real-valued functions

Let $g \in C^1[a, b]$, and suppose we want to find the solution to the equation $g(x) = 0$ on $[a, b]$. We note that we can always rewrite this equation as $x = g(x) + x$, and then out problem is equivalent with finding a fixed point of the function $f(x) = x + g(x)$.

Theorem 2.1 (Differentiable Contraction)

Let $(\mathbb{R}, d)$ be a metric space of real numbers with the metric $d(x, y) = |x - y|$ and let $[a, b]$ be a closed interval in $\mathbb{R}$. Moreover, let $f : [a, b] \to [a, b]$ be a continuous and differentiable function such that $\sup\limits_{x \in [a,b]} |f'(x)| \leq k < 1$. Then there exists a unique fixed point $x^* \in [a, b]$ of f.

Example 2.1

We want to find the solution to the equation $cos(x) - 2x = 0$ on $[0, \pi]$. Then we can write this equation as $x = \frac{1}{2} cos(x)$, and try to find the fixed point of the function $f(x) = \frac{1}{2} cos(x)$ on $[0, \pi]$. We have to show that f is a contraction on $[0, \pi]$. To do so, we apply the theorem 2.1. We have

$$\sup_{x \in [0,\pi]} |f'(x)| = \sup_{x \in [0,\pi]} \left| -\frac{1}{2} sin(x) \right| = \frac{1}{2} < 1.$$

We have shown that f is a contraction and, by the Banach's FPT, it has a fixed point x^* that is the limit of the sequence $\{x_n\}$ generated by the scheme $x_{n+1} = f(x_n)$ with any starting point $x_0 \in [0, \pi]$.
Note that to show that f is a contraction we could also directly apply the definition:

$$\begin{aligned} |f(x) - f(y)| &= \left| \frac{1}{2} cos(x) - \frac{1}{2} sin(x) \right| = \left| sin\left(\frac{x+y}{2}\right) sin\left(\frac{x-y}{2}\right) \right| \\ &\leq \sup_{x,y \in [0,\pi]} \left| sin\left(\frac{x+y}{2}\right) \frac{1}{2} |x - y| \right| = \frac{1}{2} |x - y| \leq |x - y|. \end{aligned}$$

2.2 Applications to integral equations

We consider integral equations in the following form

$$f(x) = g(x) + \mu \int_a^b k(x,y) f(y) dy,$$

where $f : [a,b] \to \mathbb{R}$ is an unknown function, $g : [a,b] \to \mathbb{R}$, and $k : [a,b]^2 \to \mathbb{R}$ are given functions, and μ is a parameter.
The above integral equation can be considered in various function spaces. Here we consider this equation only in $(C[a,b], d)$ with $d(f,g) = \sup_{x\in[a,b]} |f(x) - g(x)|$.
We assume that $g \in C[a,b]$, and that the kernel k is continuous on the square $[a,b]^2$, which implies that k is bounded on $[a,b]^2$, meaning that there exists a constant c, such that $|k(x,y)| \leq c$ for all $(x,y) \in [a,b]^2$.

Theorem 2.2

The metric space $(C[a,b], d)$ is complete

Note that our integral equation can be rewritten as $T(f) = f$, where

$$T(f)(x) = g(x) + \mu \int_a^b k(x,y) f(y) dy.$$

First we have to show that the mapping $T : C[a,b] \to C[a,b]$ is well-defined, but this is obvious, as g and k are both continuous on their domains. Let us now determine for which values of μ the map T is a contraction. We have

$$\begin{aligned} d(T(f_1), T(f_2)) = \sup_{x\in[a,b]} |T(f_1)(x) - T(f_2)(x)| = \sup_{x\in[a,b]} |\mu| \left| \int_a^b k(x,y)(f_1(y) - f_2(y)) dy \right| \leq \\ \leq |\mu| \sup_{x\in[a,b]} \int_a^b |k(x,y)||f_1(y) - f_2(y)| dy \leq c|\mu| \sup_{x\in[a,b]} |f_1(x) - f_2(x)| \int_a^b dy = \\ = c|\mu|(b-a) d(f_1, f_2). \end{aligned}$$

It is now required that $c|\mu|(b-a) < 1$, or $|\mu| < \frac{1}{c(b-a)}$, for T to be a contraction. Applying the Banach's FPT, we see that the map T has a unique fixed point $f^* \in C[a,b]$.

Theorem 2.3 (Integral equation)

Consider the integral equation

$$f(x) = g(x) + \mu \int_a^b k(x,y) f(y) dy.$$

Suppose that k and g are continuous on $[a,b]^2$ and $[a,b]$ respectively, and assume that the parameter μ satisfies $|\mu| < \frac{1}{c(b-a)}$ with the constant c such that $|k(x,y)| < c$ for all $(x,y) \in [a,b]^2$.

Then the integral equation has a unique solution $f \in C[a,b]$. Moreover, this solution is a limit of the sequence $\{f_n\}$ where f_0 is a continuous function on $[a,b]$, and

$$f_{n+1} = g(x) + \mu \int_a^b k(x,y) f_n(y) dy.$$

2.3 Applications to differential equations

Let's consider the initial value problem

$$\begin{aligned} x'(t) &= f(t, x(t)) \\ x(t_0) &= x_0 \end{aligned}$$

where $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ is a given function and $x(t)$ is an unknown function that we want to find.

Theorem 2.4 (Picard-Lindelöf)

Let f be continuous on the rectangle

$$R = \{(t,x) \in \mathbb{R}_+ \times \mathbb{R} : |t - t_0| \leq a, |x - x_0| \leq b\}$$

and thus bounded on R, say $|f(t,x)| \leq c$ for all $(t,x) \in R$. Suppose that f satisfies the Lipschitz condition on R with respect to the second argument, i.e., there exists a constant k such that

$$|f(t,x) - f(t,y)| \leq k|x - y|$$

for all $(t,x), (t,y) \in R$.
Then the initial value problem has a unique solution, which exists on the interval $[t_0 - \beta, t_0 + \beta]$, where

$$\beta = \min\left\{a, \frac{b}{c}, \frac{1}{k}\right\}.$$

Corollary 2.1 (Picard-Lindelöf)

Under the assumptions of Picard-Lindelöf theorem, the sequence given by

$$\begin{aligned} x_0(t) &= x_0 \\ x_{n+1}(t) &= T(x_n)(t) = x_0 + \int_{t_0}^t f(s, x_n(s)) ds \end{aligned}$$

converges uniformly to the unique solution $x(t)$ on $J = [t_0 - \beta, t_0 + \beta]$.

Example 2.2

Consider the differential equation

$$x'(t) = \sqrt{x(t)} + x^3(t), \quad x(1) = 2.$$

We have

$$x_1(t) = 2 + \int_1^t \left(\sqrt{2} + 2^2\right) ds = 2 + \left(\sqrt{2} + 8\right)(t-1)$$
$$x_2(t) = 2 + \int_1^t \left(\sqrt{x_1(s)} + x_1(s)^3\right) ds = *\text{hot mess}*$$

2.4 Applications to matrix equations

Suppose we want to find a solution of the matrix equation

$$Ax = B$$

where $A \in \mathbb{R}^{n \times m}, b \in \mathbb{R}^n$.
We note that this equation can be rewritten as

$$x = (I - A)x + b$$

where I is the identity matrix.
Let's define the map $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$Tx = (I - A)x + b.$$

Then the problem of solving the matrix equation $Ax = b$ is equivalent with finding a fixed point of T.
Let's define $\alpha_{ij} = \delta_{ij} - a_{ij}$ where a_{ij} are elements of the matrix A, and δ_{ij} is the Kronecker delta. Using this notation we have

$$(Tx)_i = \sum_{j=1}^{n} \alpha_{ij} x_j + b_i$$

We will show that $Ax = b$ has a unique solution if

$$\sum_{j=1}^{n} |\alpha_{ij}| \leq \alpha < 1.$$

for all $i = 1, 2, \ldots, n$. Consider the matrix space $(\mathbb{R}^n, d)$, with $d(x, y) = \max_{1 \leq i \leq n} |x_i - y_i|$ for $x, y \in \mathbb{R}^n$
. We have

$$
\begin{aligned}
d(Tx, Ty) = \max_i |(Tx)_i - (Ty)_i| = \\
&= \max_i \left| \sum_{j=1}^{n} \alpha_{ij}(x_j - y_j) \right| \leq \\
&\leq \max_i \sum_{j=1}^{n} |\alpha_{ij}||x_j - y_j| \leq \\
&\leq \max_i \sum_{j=1}^{n} |\alpha_{ij}| \cdot \max_j |x_j - y_j| = \\
&= \max_i \sum_{j=1}^{n} |\alpha_{ij}| \cdot d(x, y).
\end{aligned}
$$

We notice that if $\sum_{j=1}^{n} |\alpha_{ij}| < 1$, for all $i = 1, 2, \ldots, n$, then $\max_i \sum_{j=1}^{n} |\alpha_{ij}| < 1$. We have

$$
\sum_{j=1}^{n} |\alpha_{ij}| = |a_{i1}| + |a_{i2}| + \cdots + |1 - a_{ii}| + \cdots + |a_{in}| < 1,
$$

so

$$
\sum_{j=1, j \neq i}^{n} |a_{ij}| < 1 - |1 - a_{ii}| < |a_{ii}|.
$$

We get the condition for the matrix A for which T is a contraction]. This condition is given by

$$
|a_{ii}| > \sum_{j=1, j \neq i}^{n} |a_{ij}|,
$$

or, in other words, matrix A should be strictly diagonally dominant.

Theorem 2.5 (Matrix equation)

The matrix equation $Ax = b$ with $A \in \mathbb{R}^{n \times n}$ an $b \in \mathbb{R}^n$ has a unique solution $x \in \mathbb{R}^n$ if A is strictly diagonally dominant. The iteration method is as follows

$$
x_{n+1} = (I - A)x_n + b, \quad x_0 \in \mathbb{R}^n.
$$

In general, we can rewrite the equation $Ax = b$ as $Qx = (Q - A)x + b$, where $Q \in \mathbb{R}^{n \times n}$. We then have the following iterative scheme

$$
Qx_{n+1} = (Q - A)x_n + b.
$$

Examples:

- $Q = I$ - Richardson method,
- Q diagonal, with $q_{ii} = a_{ii}$ - Jacobi method,
- $Q = D - L$ with D diagonal, and L lower triangular - Gauss-Seidel method.

3. Normed spaces

Def 3.1 (Norm and normed space)

A norm on a vector space X is a real-valued function denoted by $||\cdot||$ which satisfies the following conditions:

- $||x|| \geq 0$ for all x. $||x|| = 0$ iff $x = 0$,
- $||\alpha x|| = |\alpha| ||x||$ for any α, and $x \in X$,
- $||x+y|| \leq ||x|| + ||y||$ for all $x, y \in X$.
  A normed space is a vector space equipped with a norm, depicted by $(X, ||\cdot||)$, or with a shorthand X.

Remark 3.1.1

A norm on X defines the metric $d(\cdot,\cdot)$ on $X \times X$, which is defined by $d(x,y) = ||x-y||$, and is called the metric induced by the norm $||\cdot||$.

Remark 3.1.2

Every normed space X is a metric space, converse might not be true.
For example, a metric defined by

$$d(x,y) = \begin{cases} 1, & x = y, \\ 0, & x \neq y, \end{cases}$$

then $||\alpha(x-y)|| = d(\alpha x, \alpha y) \neq |\alpha| d(x,y) = |\alpha| ||x-y||$.

Lemma 3.1 (Norm continuity)

The norm $||\cdot||$ defined on X is a continuous mapping of X into $\mathbb{R}$.

Examples of normed spaces

- $(\mathbb{R}^n, ||\cdot||_2)$, with $||x||_2 = \left(\sum_{m=1}^n x_i^2\right)^{\frac{1}{2}}$,
- $(C[a,b], ||\cdot||)$, with $||f|| = \max\limits_{x\in[a,b]} |f(x)|$,
- $(L^p(\Omega), ||\cdot||_{L^p(\Omega)})$, with $\Omega \subset \mathbb{R}$, $p \geq 1$ and

$$||f||_{L^p(\Omega)} = \begin{cases} \left(\int_\Omega f(x)^p dx\right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ ess \sup\limits_{x\in\Omega} |f(x)|, & p = \infty. \end{cases}$$

Def 3.2 (Norm equivalence)

Two normed spaces $(X, ||\cdot||_1)$, $(X, ||\cdot||_2)$ are called topologically equivalent, or two norms $||\cdot||_1$, and $||\cdot||_2$ are called equivalent if there exist positive constants C_1, and C_2, such that

$$C_1||x||_2 \leq ||x||_1 \leq C_2||x||_2$$

for all $x \in X$.

Theorem 3.1 (Equivalence of norms in finite dimensional spaces)

All norms of finite dimensional space X are equivalent.

Def 3.3 (Convergence in normal spaces)

A sequence $\{x_n\}$ in a normed space $(X, ||\cdot||)$ is convergent if there exists $x \in X$ such that $\lim\limits_{n\to\infty} ||x_n - x|| = 0$.

Def 3.4 (Cauchy sequence)

A sequence $\{x_n\}$ in a normed space $(X, ||\cdot||)$ is a Cauchy sequence if

$$\lim_{m,n\to\infty} ||x_n - x_m|| = 0.$$

Def 3.5 (Complete space)

We say that a normal space $(X, ||\cdot||)$ is complete if every Cauchy sequence $\{x_n\}$ in X is convergent to some $x \in X$.

Def 3.6 (Banach space)

A complete normed space is called a Banach space.

Theorem 3.2 (Euclidean space is complete)

The space $(\mathbb{R}^N, ||\cdot||_2)$ with the standard Euclidean norm is complete.

Theorem 3.3 ()

Let Ω be a compact subset of $\mathbb{R}^n$. Then the set $C(\Omega)$ of all continuous functions on Ω equipped with the norm $||f|| = \max\limits_{x\in\Omega} |f(x)|$ is a Banach space.

4. Hilbert spaces

Def 4.1 (Inner product and inner product space)

Let X be a vector space over the field $\mathbb{F}$ over the real or complex numbers. A mapping $\langle\cdot,\cdot\rangle : X^2 \to \mathbb{F}$ is called an inner product if for all $x, y \in X$ the following conditions are satisfied

1. $\langle x, x\rangle \geq 0$, and $\langle x, x\rangle = 0 \iff x = 0$,
2. $\langle x, y\rangle = \overline{\langle y, x\rangle}$,
3. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ for $\alpha \in \mathbb{F}$,
4. $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$.
   The vector space X together with an inner product $\langle\cdot,\cdot\rangle$ is called an inner product space or pre-Hilbert space and is denoted $(X, \langle\cdot,\cdot\rangle)$.

Remark 4.1

- $\overline{\langle x, y\rangle}$ denotes the complex conjugate of $\langle x, y\rangle$,
- The condition 2 implies that $\langle x, x\rangle$ must be a real number,
- If $\mathbb{F} = \mathbb{R}$ then $\langle x, y\rangle = \langle y, x\rangle$,
- Conditions 3 and 4 imply that the function $\langle\cdot,\cdot\rangle$ is linear in the first variable. It is easy to see that $\langle\cdot,\cdot\rangle$ is also linear in the second variable if $\mathbb{F} = \mathbb{R}$,

Examples

- $\mathbb{R}^N$, with $\langle x, y\rangle = \sum_{i=1}^N x_i y_i$,
- $C(\Omega)$, with $\langle f, g\rangle = \int_\Omega f(x)\overline{g(x)}dx$,
- $L_2(\Omega)$, with $\langle f, g\rangle = \int_\Omega f(x)g(x)dx$.

Theorem 4.1 (Cauchy-Schwarz-Bunyakowski inequality)

For all $x, y \in (X, \langle\cdot,\cdot\rangle)$ we have

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle.$$

Theorem 4.2 (Inner product space is normed)

Every inner product space $(X, \langle\cdot,\cdot\rangle)$ is a normed space with respect to the norm $||x|| = \sqrt{|\langle x, x\rangle|}$
.

Def 4.2 (Hilbert space)

An inner product space $(X, \langle\cdot,\cdot\rangle)$ is called a Hilbert space if the normed space $(X, ||\cdot||)$ with the norm induced by the inner product is a Banach space.

Theorem 4.2 (Parallelogram law)

Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space. Then for all $x, y \in X$ we have

$$||x+y||^2+||x-y||^2=2||x||^2+2||y||^2.$$

Remark 4.2

The parallelogram law is not valid for an arbitrary norm on a vector space.

Theorem 4.4 (Polarisation identity)

For any two elements x, y in an inner product space we have

$$\langle x,y\rangle=\frac{1}{4}\left(||x+y||^2-||x-y||^2+i||x+iy||^2-i||x-iy||^2\right).$$

Theorem 4.5 (Normed space is inner product space sometimes)

A normed space is an inner product space if and only if the norm of the normed space satisfies the parallelogram law.

5. Linear operators

Def 5.1 (Linear operator)

Let $(X, ||\cdot||_X)$ and $(Y, ||\cdot||_Y)$ be Banach spaces, and let $A: X\to Y$ be a map. We say that A is linear if $A(\alpha x+\beta y)=\alpha Ax+\beta By$ for all $x, y\in X$, and $\alpha, \beta\in\mathbb{R}$.

Def 5.2 (Bounded operator)

A linear operator is bounded if there exists a constant $M>0$ such that

$$||Ax||_Y\leq M||x||_X$$

for all $x\in X$.
We denote the set of all linear and bounded operators as

$$\mathcal{L}(X,Y)=\{\ A: X\to Y: A \text{ is linear and bounded}\}.$$

Def 5.3 (Operator norm)

A set $\mathcal{L}(X,Y)$ can be equipped with the operator norm

$$||A||_{op}=\inf\{M\,|\,||Ax||_Y\leq M||x||_X\}=\sup_{x\neq 0}\frac{||Ax||_Y}{||x||_X}=\sup_{||x||_X=1}||Ax||_Y.$$

Theorem 5.1 (Operator set is a Banach space)

The set $\mathcal{L}(X,Y)$ equipped with $||\cdot||_{op}$ norm is a Banach space.

Theorem 5.2 (Bounded iff continuous)

Let $A : X \to Y$ be a linear operator, then A is bounded if and only if A is continuous.

Example 5.1

Consider a matrix $A \in \mathbb{R}^{m\times n}$. The matrix A is a linear operator from $(\mathbb{R}^n, ||\cdot||_\alpha)$ to $(\mathbb{R}^m, ||\cdot||_\beta)$ and the corresponding induced norm (or the operator norm) on the space $\mathbb{R}^{m\times n}$ is defined by

$$||A||_{op} = \sup\{||Ax||_\beta \,:\, ||x||_\alpha = 1\}.$$

Example 5.2

Let $\Omega \subset \mathbb{R}^n$ be an open and bounded set. We consider the integral operator $K : L^2(\Omega) \to L^2(\Omega)$ defined as follows

$$Ku(x) = \int_\Omega k(x,y)u(y)dy, \text{ with } \iint_{\Omega^2} |k(x,y)|^2 dxdy = c < \infty.$$

It can be shown that the operator K is bounded.

Def 5.4 (Unbounded linear operator)

An unbounded linear operator $A : X \to Y$ is a pair $(A, D(A))$, where $D(A)$ is a linear subspace of X and A is not not bounded on $D(A)$.

Example 5.3

Consider $A = -\frac{d^2}{dx^2}$ on $L^2(\Omega)$. Since $C^2(\Omega) \subset L^2(\Omega)$, we define the operator A only on its domain.

$$A : \left\{f \in C^2(\Omega) : Af \in L^2(\Omega)\right\} \to Y.$$

Let $f(x) = e^{-kx}$. Then

$$||Af(x)||^2 = \int_0^1 \left(-\frac{d^2}{dx^2}e^{-kx}\right)dx = k^4\int_0^1 e^{-2kx}dx = \frac{k^3}{2}(1-e^{-2k}).$$

Def 5.5 (Operator range)

The range of operator $A : D(A) \to Y$ is defined as

$$R(A) = \{g \in Y : g = Af, f \in D(A)\}.$$

The kernel (null space) of the operator $A : D(A) \to Y$ is defined as

$$Ker(A) = N(A) = \{f \in D(A) : Af = 0\}.$$

Theorem 5.3 (Invertible operator)

The linear operator $A : D(A) \to Y$ is invertible if an only if $Ker(A) = \{0\}$.

Def 5.6 (Operator bounded from below)

We say that a linear operator $A : X \to Y$ is bounded from below if there exists constant $C > 0$ such that

$$||Ax||_Y \geq C||x||_X.$$

Theorem 5.4 (Bounded operator is invertible)

Let $A : X \to Y$ be a linear operator. Then the following propositions are equivalent

- A is bounded from below,
- $A^{-1} : R(A) \to X$ exists and is bounded.

7. Introduction to inverse problems

Def 7.1 (Inverse problem)

An inverse problem is the task of recovering the parameter $u \in X$ from measured data $f \in Y$, when $f = Au + e$. Here

- X and Y are vector spaces with appropriate topologies, whose elements represent model parameters and data, respectively,
- $A : X \to Y$ (forward operator) is a known, continuous operator, that maps model parameters to data in absence of noise,
- $e \in Y$ is a sample of random variable modelling the observation noise.
  Inverse problems are usually ill-posed.

Def 7.2 (Well-posed inverse problem)

The inverse problem is well posed if the following three conditions hold:

- It has a solution (existence),
- The solution is unique (uniqueness),
- The solution depends continuously on the data (stability).
  If at least one of the conditions fails, we say that the inverse problem is ill-posed.

7.1 Variational methods

The idea behind the variational methods

TODO: last lecture

9. The Sobolev spaces

Def 9.1 (Sobolev space)

Let $\Omega \subset \mathbb{R}^n$ be an open set. The Sobolev space $H^k(\Omega)$ is defined by

$$H^k(\Omega) = \big\{u \in L^1_{loc}(\Omega) \mid D^\alpha u \in L^2(\Omega),\ 0 \leq |\alpha| \leq k\big\}.$$

with the norm

$$||u||^2_{H^k(\Omega)} = \sum_{0\leq|\alpha|\leq k} ||D^\alpha u||^2_{L^2(\Omega)}.$$

The Sobolev space $H^k(\Omega)$ is the Hilbert space with the inner product

$$\langle u, v\rangle_{H^k(\Omega)} = \sum_{0\leq|\alpha|\leq k} \langle D^\alpha u, D^\alpha v\rangle.$$

In the particular case $\Omega \subset \mathbb{R}$ and $k = 1$, we have

$$H^1(\Omega) = \big\{u \in L^1_{loc}(\Omega) \mid u \in L^2(\Omega)\ ,\ u' \in L^2(\Omega)\big\},$$

$$||u||^2_{H^1(\Omega)} = \int_\Omega \big(u^2 + (u')^2\big)dx,$$

$$\langle u, v\rangle_{H^1(\Omega)} = \int_\Omega \big(uv + u'v'\big)dx,$$

where u', v' are the weak derivatives of u and v.
By $H^1_0(\Omega)$ we denote the subspace of $H^1(\Omega)$ given by

$$H^1_0(\Omega) = \big\{u \in H^1(\Omega) \mid u = 0 \text{ on } \partial\Omega\big\}.$$

Theorem 9.1 (The Poincaré inequality)

Let $1 \leq k \leq \infty$ and $\Omega \subset \mathbb{R}^n$ bounded at least in one direction. Then there exists a constant $c > 0$ dependent only on k and Ω such that for every function $u \in H^k_0(\Omega)$ we have

$$||u||_{L^k(\Omega)} \leq c||\nabla u||_{L^k(\Omega)}.$$

Theorem 9.2 (The Lax-Milgram theorem)

Let V be a vector space with the inner product $\langle\cdot,\cdot\rangle$, and the associated norm $||\cdot|| = \sqrt{\langle\cdot,\cdot\rangle}$. Let $a : V^2 \to \mathbb{R}$ be a bilinear form which satisfies the following:

- It is continuous, i.e. there exists $M > 0$ such that $|a(u, v)| \leq M||u||||v||$ for all $u, v \in V$,
- It is coercive, i.e. there exists $\beta > 0$ such that $|a(v, v)| \geq \beta||v||^2$ for all $v \in V$.
  Then for any linear continuous form l on V there exists a unique $u \in V$ such that

$$a(u, v) = l(v), \text{ for all } v \in V.$$

Example 9.1 (Dirichlet problem)

Let $\Omega \subset \mathbb{R}^n$ ad $f : \Omega \to \mathbb{R}$ be a given function. We consider the Dirichlet boundary value problem for the Laplace operator

$$\begin{aligned} -\Delta u &= f, \text{ in } \Omega, \\ u &= 0, \text{ on } \partial\Omega. \end{aligned}$$

Theorem 9.3 (Weak formulation of Dirichlet problem)

We have:

- For every $f \in L^2(\Omega)$ there exists a unique $u \in H_0^1(\Omega)$ which satisfies:

$$\begin{aligned} \int \nabla u \nabla v dx &= \int f v dx, \\ u &\in H_0^1(\Omega), \end{aligned}$$

  for every $v \in H_0^1(\Omega)$.
- The solution u of the Dirichlet problem satisfies:

$$\begin{aligned} -\Delta u &= f, \text{ in } \mathcal{D}'(\Omega) \text{ (equality as distributions)}, \\ u &= 0, \text{ on } \partial\Omega. \end{aligned}$$

- For $u \in H_0^1(\Omega)$ the two above problems are equivalent.

Def 9.2 (Properties of bilinear forms)

Let $a : V^2 \to \mathbb{R}$ be a bilinear form. We say that:

- a is symmetric if $a(u, v) = a(v, u)$ for all $u, v \in V$,
- a is positive if $a(v, v) \geq 0$ for all $v \in V$,
- a is positive definite if it is positive and $a(v, v) = 0$ holds only for $v = 0$.

Theorem 9.4 (Solution to Dirichlet problem)

Let V be a vector space, $l : V \to \mathbb{R}$ be a linear form, and $a : V^2 \to \mathbb{R}$ be a bilinear, symmetric, positive form. Then the following statements are equivalent

- $u \in V$ is a unique solution to $a(u,v) = l(v)$ for every $v \in V$,
- there exists a $u \in V$ such that $J(u) \leq J(v)$ for all $v \in V$, where $J(v) = \frac{1}{2}a(v,v) - l(v)$.

Corollary

The weak solution u of the Dirichlet problem is a solution of the minimisation problem

$$\begin{cases} J(u) \leq J(v), \quad \forall v \in H_0^1(\Omega), \\ u \in H_0^1(\Omega). \end{cases}$$

Theorem 9.5 (Unique solution to minimisation problem)

Let V be a linear space and $J : V \to \mathbb{R}$ be a strictly convex functional. Then there exists at most one solution u to the minimisation problem.

$$J(u) \leq J(v), \quad \forall v \in V, \quad u \in V.$$

Theorem 9.6 (Convex bilinear form)

Let V be a linear vector space and $a : V^2 \to \mathbb{R}$ be a bilinear form which is symmetric, and positive definite. Then the quadratic form $q : V \to \mathbb{R}$, which is defined by $q(v) = a(v,v)$, is strictly convex.

10. Digression

Consider the following problem

$$\begin{cases} -\left(a(x)u'(x) = f(x)\right), \\ u'(0) = u'(1) = 0, \end{cases}$$

where we are given f, and u, and are searching for a. We want to decide if the above problem is well-posed, as in if the solution exists, is unique, and depends continuously on the given conditions. We shall derive the required conditions that u should fulfil to satisfy

$$||a_1 - a_2|| \leq C||u_1 - u_2||.$$

Let u_j denote the solution for a_j, with $j = 1, 2$. We have

$$a_1(x) - a_2(x) = -\frac{1}{u_1'(x)u_2'(x)} \int_0^x f(y)dy \left(u_2'(x) - u_1'(x)\right),$$

squaring both sides, and integrating, we get

$$||a_1 - a_2||^2 \leq \frac{1}{\gamma^4}\left(\int_0^1 f(x)dx\right)^2 \int_0^1 \left(u_2'(x) - u_1'(x)\right)^2 dx.$$

So here we assumed that

$$u \in C^2([0,1]), \quad 0 < \gamma < u'(x), \quad ||u||_{C^2} \leq M.$$

Integration by parts and the Cauchy-Schwarz inequality yields

$$\begin{aligned}\int_0^1 \left(u_2'(x) - u_1'(x)\right)^2 dx = \int_0^1 (u_1(x) - u_2(x)) \left(u_2''(x) - u_1''(x)\right) dx \leq \\ \leq ||u_1 - u_2||_{L^2(0,1}||u_1'' - u_2''||_{L^2(0,1)} \leq \\ \leq 2M||u_1 - u_2||_{L^2(0,1)}.\end{aligned}$$

So we have

$$||a_1 - a_2||_{L^2(0,1)} \leq \sqrt{2M}||u_1 - u_2||_{L^2(0,1)},$$

and the space that u should belong to is

$$C = \left\{u \in C^2([0,1]) : \quad 0 < \gamma < u'(x), \quad ||u||_{C^2} \leq M\right\}.$$

If u is measured in a real world, we would like to regularise it, to fit into the space C.

Regularisation and iterative reconstruction

Consider the following problem

$$\begin{cases} -\nabla(a\nabla u) = f, \\ \frac{\partial u}{\partial n} = 0. \end{cases}$$

Assume that measurements u_δ of the exact solution u for $x \in \Omega$ are given. We define the objective functional

$$J(u,a) = \int_\Omega (u(x) - u_\delta(x))^2 dx,$$

where u implicitly depends on a, via the differential equation at hand. Also, $J : H^1(\Omega)^2 \to \mathbb{R}$.

We denote by $u_a \in H^1(\Omega)$ the unique weak solution to the PDE. Moreover, we impose on a to be a function in $H^1(\Omega)$.

We introduce the so-called *reduced objective functional*:

$$J'(a) = J(u_a, a) + \frac{\alpha}{2}||a||^2_{H^1(\Omega)}.$$

Then we get the constrained optimisation problem

$$\min_{a \in H^1(\Omega)} J'(a) \qquad (*)$$

subject to

$$\begin{cases} -\nabla(a\nabla u) = f, \\ \frac{\partial u}{\partial n} = 0. \end{cases}$$

To solve this problem we introduce the Lagrangian

$$\mathcal{L}(u,a,p) = J(u,a) + \frac{\alpha}{2}||a||^2_{H^1(\Omega)} + \int_\Omega a\nabla u \nabla p dx - \int_\Omega fp dx,$$

where $p \in H^1(\Omega)$ is a Lagrange multiplier or the adjoint variable.

The method of Lagrange multipliers states that te solution to the problem $(*)$ has to be a stationary point of the Lagrangian, that is it has to satisfy the following set of equations:

$$\begin{cases} \delta_u \mathcal{L}(u,a,p,h_u) &= 0, \quad \forall h_u \in H^1(\Omega), \\ \delta_a \mathcal{L}(u,a,p,h_a) &= 0, \quad \forall h_a \in H^1(\Omega), \\ \delta_p \mathcal{L}(u,a,p,h_p) &= 0, \quad \forall h_p \in H^1(\Omega), \end{cases}$$

where for $F : \mathcal{U} \to \mathbb{R}$, $\delta_u F(u,h)$ is defined by

$$\lim_{\epsilon \to 0} F(u,h) = \lim_{\epsilon \to 0} \frac{F(u+\epsilon h) - F(u)}{\epsilon}.$$

Then we get the following:

$$\begin{cases} -\nabla(a\nabla p) &= -(u-u_\delta), & \frac{\partial u}{\partial n} = 0, & (1) \\ \alpha a - \alpha \Delta a &= -\nabla u \nabla p & \frac{\partial a}{\partial n} = 0, & (2) \\ -\nabla(a \nabla u) &= f, & \frac{\partial p}{\partial n} = 0. & (3) \end{cases}$$

The algorithm (The Landweber iterative method)

Input: u_δ, $\tau > 0$, $\alpha > 0$.
Output: u, a.

Set $j = 0$, initialise $a_0(x) \in H^1(\Omega)$
Repeat until satisfied:

1. Solve (3) for u_j with a_j,
2. Solve (1) for p_j with a_j and u_j,
3. Solve (2) for a with a_{j-1}, u_j, and p_j,
4. Update $a_{j+1} = a_j - \tau a$,
5. Set $j = j + 1$.

11. Vector space bases

Def 11.1 (Orthogonal and orthonormal sets)

We call a set of vectors $\{e_i\}$ in a vector space with a scalar product orthogonal when

$$\langle e_i, e_j \rangle = 0, \quad i \neq j.$$

Additionally if $||e_i||^2 = 1$ we call it orthonormal, in short $\langle e_i, e_j \rangle = \delta_{ij}$.

Fact

The orthogonal set $\{e_i\}$ is linearly independent.

Theorem 11.2 (Bessel inequality)

Let $\{e_i\}$ be an orthonormal set in a Hilbert space $\mathcal{H}$, and ψ any vector in $\mathcal{H}$. Then

$$\sum_{i \in I} |\langle \psi, e_i \rangle|^2 \leq ||\psi||^2.$$

Theorem 11.3

Let $\{e_i\}$ be an orthonormal set of vectors in a Hilbert space $\mathcal{H}$. Then:

- $\sum_{n=1}^{\infty} \alpha_i e_i$ converges iff $\sum_{n=1}^{\infty} |\alpha_i|^2$ converges,
- If $\psi = \sum_{n=1}^{\infty} \alpha_i e_i$ and $\psi = \sum_{n=1}^{\infty} \beta_i e_i$ then $a_i = \beta_i$ for all i.

Theorem 11.4

The following conditions are equivalent for a Hilbert space $\mathcal{H}$ and an orthonormal set $\{e_i\}$:

- If $\langle \psi, e_i \rangle = 0$ for all $i \in I$, then $\psi = 0$,
- Any $\psi \in \mathcal{H}$ has a form $\psi = \sum_{i \in I} \langle \psi, e_i \rangle e_i$,
- For any $\psi \in \mathcal{H}$, $||\psi||^2 = \sum_{i \in I} |\langle \psi, e_i \rangle|^2$.
  We call such set $\{e_i\}$ a complete orthonormal set or just a basis of $\mathcal{H}$.

Remarks

- Alas, bases can be uncountable, for example:
  $l^2(\mathbb{R})$ - all functions that are non-zero at finite set of points in $\mathbb{R}$ such that
  $$\sum_{x \in \mathbb{R}} |f(x)|^2 < \infty.$$
- Nearly always we deal with countable bases. Such Hilbert spaces are called separable.

Theorem 11.5

Given a basis there exists an isomorphism between separable $\mathcal{H}$, and l^2.

Examples