Chapter II
Unconstrained and constrained optimization

Skills to acquire

- **Fundamental results about existence of minimizers.**
- **Lagrange multipliers for constrained optimization problems.**
- **Least squares problems.**

1 Generalities

Let U be an open subset of $\mathbb{R}^n$ and $f : U \subset \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$. We aim at minimizing f, i.e., finding some $v \in U$ such that

$$v = \underset{u \in U}{\operatorname{argmin}}\ f(x). \tag{II.1}$$

Definition 1 (Sublevel and epigraph)**.** A c-sublevel set of f is:

$$S_c f = \{x \in \operatorname{dom} f : f(x) \leq c\}.$$

The epigraph of f is defined by

$$\operatorname{epi} f = \{(x, y) \in \mathbb{R}^{n+1} : f(x) \leq y\}.$$

f is said to be closed if its epigraph is closed.

Proposition 1. f is closed if and only if all its sublevel sets are closed.

Proposition 2. If f is closed with bounded sublevel sets then it has a minimizer.

Semi-continuity is a property of extended real-valued functions that is weaker than continuity. A function f is upper (respectively, lower) semi-continuous at a point x_0 if, roughly speaking, the function values for arguments near x_0 are not much higher (respectively, lower) than $f(x_0)$.

Definition 2 (Semi-continuity)**.** f is lower semi-continuous (l.s.c.) at $x_0 \in \operatorname{dom} f$ if

$$\forall \epsilon > 0, \exists \delta > 0 : f(x) \geq f(x_0) - \epsilon, \forall x \in B_\delta(x_0)$$

or, for metric spaces

$$f(x_0) \leq \liminf_{x \to x_0} f(x).$$

The function f is l.s.c. if it is for every $x_0 \in \operatorname{dom} f$.

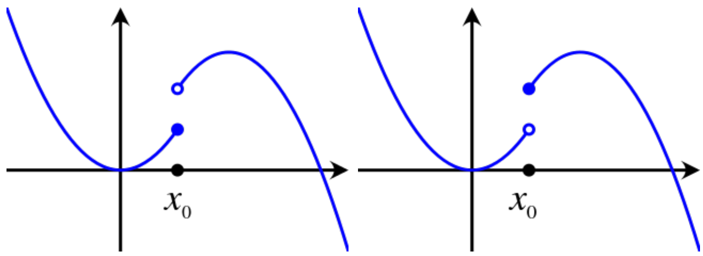

Fig. II.1: Lower and upper semi-continuous

Theorem 1. f is closed $\iff$ f is l.s.c.

Theorem 2 (Weierstrass theorem for closed functions)**.** Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper closed function and assume that C is a compact set satisfying $C \cap \operatorname{dom}(f) \neq \emptyset$. Then f is bounded below over C and attains its minimal value over C.

The compactness of C is replaced by closedness if the function has a property called coerciveness.

Definition 3 (coercivity)**.** A proper function $f : \mathbb{R}^n \to (-\infty, +\infty]$ is called coercive if

$$\lim_{\|x\| \to +\infty} f(x) = +\infty.$$

Theorem 3 (attainment under coercivity)**.** Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper closed and coercive function and let $S \subseteq \mathbb{R}^n$ be a nonempty closed set satisfying $S \cap \operatorname{dom}(f) \neq \emptyset$. Then f attains its minimal value over S.

2 Unconstrained problem

Consider the following optimization problem:

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x}) \tag{II.2}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is sufficiently differentiable.

The next two theorems give fundamental results of existence and characterization of minimizers.

Theorem 4 (First-order necessary conditions)**.** If $\bar{\boldsymbol{x}}$ is a local minimizer of f and f is continuously differentiable in an open neighborhood of $\bar{\boldsymbol{x}}$, then $\nabla f(\bar{\boldsymbol{x}}) = 0$.

Theorem 5 (Second-order conditions)**.** Suppose that $\nabla^2 f$ is continuous in an open neighborhood of $\bar{\boldsymbol{x}}$. If $\bar{\boldsymbol{x}}$ is a local minimizer of f then $\nabla f(\bar{\boldsymbol{x}}) = 0$ and $\nabla^2 f$ is positive semi-definite. Conversely, if $f(\bar{\boldsymbol{x}}) = 0$ and $\nabla^2 f$ is positive definite, then $\bar{\boldsymbol{x}}$ is a strict local minimizer of f.

Example 1 (Mean and median)**.** Consider $(x_i)_{i=1,\ldots,n} \in \mathbb{R}$. We would like to compute the best value $\bar{x} \in \mathbb{R}$ which approximates the whole set, i.e., which minimizes the following function:

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} (x - x_i)^2.$$

This criterion is called the *Mean Squared Error* (MSE). From the necessary and sufficient conditions, one can easily show that the solution is the arithmetic mean

$$\boxed{\bar{x} = \frac{x_i}{n}}.$$

If we consider the *Mean Absolute Error (MAE)*:

$$f(x) = \sum_{i=1}^{n} |x - x_i|\,,$$

and restrict the search space to the set $\{x_i, ..., x_i\}$, we can prove that the solution is the median value:

$$\boxed{\bar{x} = \text{median}\{x_1, \cdots, x_n\}}.$$

Indeed, let i such that $x \in [x_i, x_{i+1}]$. We have:

$$f(x) = \sum_{j \leqslant i} (x - x_j) + \sum_{j > i} (x_j - x)$$

which is a linear function (when restricted to the interval $[x_i, x_{i+1}]$; with gradient equal to the number of j such that $j \leqslant i$ minus the the number of j such that $j > i$. Thus the gradient is zero if it is less than a median, it is positive if greater than a medium, and zero precisely when there are as many data points to the left and right. In other words, x decreases as we approach a medium from below, stays constant on the set of mediums, and then decreases afterwards. So minimums occurs exactly at mediums.

Example 2 (Linear regression)**.** Consider $(x_i, y_i)_{i=1,\ldots,n} \in \mathbb{R}^2$. We would like to approximate this point cloud by an affine function $y = ax + b$, i.e., we seek $(a, b) \in \mathbb{R}^2$ which minimizes the following function:

$$J(a, b) = \sum_{i=1}^{n} (ax_i + b - y_i)^2.$$

First order conditions

$$\begin{cases} \frac{\partial J}{\partial a} & = 2\sum_{i=1}^{n} (ax_i + b - y_i)x_i = 0, \\ \frac{\partial J}{\partial b} & = 2\sum_{i=1}^{n} (ax_i + b - y_i) = 0. \end{cases}$$

We can write

$$\begin{cases} a\sum_{i=1}^n x_i^2 \ + b\sum_{i=1}^n x_i \ - \ \sum_{i=1}^n x_i y_i & = 0 \\ a\sum_{i=1}^n x_i \ + nb \ - \ \sum_{i=1}^n y_i & = 0 \end{cases}$$

We use the following notations:

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}, \quad s_x = \sum_{i=1}^n x_i, \quad s_{xx} = \sum_{i=1}^n x_i^2, \quad s_{xy} = \sum_{i=1}^n x_i y_i, \quad s_y = \sum_{i=1}^n y_i,$$

So, we obtain

$$\begin{cases} s_{xx}\, a \ + \ s_x\, b \ - \ s_{xy} & = 0 \\ s_x\, a \ + \ n\, b \ - \ s_y & = 0 \end{cases}$$

$$\boxed{b = \frac{s_x\, s_{xy} - s_{xx}\, s_y}{s_x^2 - n s_{xy}}, \qquad a = \frac{s_y}{s_x} - \frac{n}{s_x}\, b.}$$

Second order conditions

$$J''(a,b) = \begin{bmatrix} s_{xx} & s_x \\ s_x & n \end{bmatrix}$$

We note that $s_{xx} > 0$ and, thanks to the Cauchy-Schwarz inequality

$$\left(\sum_{i=1}^n u_i v_i\right)^2 \leqslant \left(\sum_{i=1}^n u_i^2\right)\left(\sum_{i=1}^n v_i^2\right),$$

with $u = (x_1, ..., x_n)$ and $v = (1, ..., 1)$, we get:

$$\left(\sum_{i=1}^n x_i\right)^2 \leqslant \left(\sum_{i=1}^n x_i^2\right) n, \text{ i.e., } s_x^2 \leq n\, s_{xx}.$$

From the previous equations, we can deduce that $J''(x)$ is positive definite.

Example 3 (Linear regression)**.** We consider the linear regression problem

$$z = \beta_0 + \beta_1 x + \beta_2 x^2,$$

i.e., we have a dataset $(x_i, y_i)_{i=1,\dots,n} \in \mathbb{R}^2$ and we want to estimate the parameters β_0, β_1 and β_2. We put

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \\ 1 & x_n & x_n^2 \end{bmatrix}; \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \ ; \ \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

The loss criterion J can be written as

$$\begin{aligned} J(\beta) &= \sum_{i=1}^n (y_i - \beta_2 x_i^2 - \beta_1 x_i - \beta_0)^2, \\ &= \|X\beta - \mathbf{y}\|^2. \end{aligned}$$

We write the first and second order conditions:

$$\frac{\partial J}{\partial \beta} = 2(X^T X)\beta - 2X^T\mathbf{y} = 0, \quad \text{and} \quad J''(\beta) = 2(X^T X).$$

So, the minimizer is the solution of the linear system

$$\boxed{X^T X\ \beta \ = \ X^T \mathbf{y}}.$$

3 Equality constrained problem

Consider the following equality constrained problem:

$$\begin{aligned} \underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}), \\ \text{subject to} \quad & h_i(\boldsymbol{x}) = 0,\, i = 1, \ldots, m. \end{aligned} \tag{II.3}$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$ are m constraints. We assume that h_i once or twice differentiable and put $\boldsymbol{h} = (h_1, \cdots, h_m)$.

To solve this problem, we can use the method of *Lagrange Multipliers*, which puts the cost function as well as the constraints in a single minimization problem:

$$\boldsymbol{x}^\star = \underset{\boldsymbol{x}}{\operatorname{argmin}} \underbrace{f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i h_i(\boldsymbol{x})}_{\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda})}. \tag{II.4}$$

where $\mathcal{L}(\mathbf{x}, \lambda)$ is called the *Lagrangian* and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ is a vector of *multipliers*.

Definition 4. A point $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ is called *feasible* when it satisfies all the constraints, i.e. $\boldsymbol{h}(\bar{\boldsymbol{x}}) = 0$. We say that a feasible point $\boldsymbol{x}$ is *regular* when $\nabla h_1(\boldsymbol{x})$, ..., $\nabla h_m(\boldsymbol{x})$ are linearly independent.

The next theorem gives necessary conditions for the minimization problem (II.3).

Theorem 6 (First-Order Necessary Conditions)**.** If a **regular** point $\boldsymbol{x}$ is a local minimizer of (II.3), then there exists $\boldsymbol{\lambda} \in \mathbb{R}^m$ such that

$$\begin{cases} \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \lambda) = \nabla_{\boldsymbol{x}} f(\boldsymbol{x}) + \sum_i \lambda_i \nabla_{\boldsymbol{x}} h_i(\boldsymbol{x}) = 0, \\ \frac{\partial \mathcal{L}(\boldsymbol{x},\lambda)}{\partial \lambda_i} \quad = \quad h_i(\boldsymbol{x}) \quad = 0,\, i = 1, \ldots, m. \end{cases} \tag{II.5}$$

Solutions of the optimality conditions are called *stationary or critical points*. Solving the optimality conditions gives also the Lagrange multipliers associated with the critical points.

> **Remark:** a local minimizer which is not a regular point might not fulfill the above optimality conditions.

Theorem 7 (Second-Order Conditions)**.** Suppose that $\boldsymbol{x}$ is a stationary point of (II.3). The, if $\boldsymbol{x}$ is a local minimizer, then Hessian of the Lagrangian, $H = \nabla^2_{\boldsymbol{x}\boldsymbol{x}}(\mathcal{L}(\boldsymbol{x}, \lambda))$, is positive semidefinite on the tangent space:

$$T = \left\{ \boldsymbol{y} \;\middle|\; \nabla h_j(\boldsymbol{x}) \boldsymbol{y} = 0, j = 1, \cdots, m \right\}.$$

Conversely, it H is positive definite on T and the first order conditions are satisfied, the $\boldsymbol{x}$ is a local minimizer.

Example 4. Consider the problem

$$\begin{aligned} \underset{(x,y) \in \mathbb{R}^2}{\text{minimize}} \quad & f(x, y) = 5x - 3y \\ \text{subject to} \quad & x^2 + y^2 = 136 \end{aligned}$$

The optimality conditions yield:

$$\begin{cases} 5 & = 2\lambda x, \\ -3 & = 2\lambda y, \\ x^2 + y^2 & = 136. \end{cases}$$

This system has two solutions: $(x_1, y_1, \lambda_1) = (-10, 6, -\frac{1}{4})$ and $(x_2, y_2, \lambda_2) = (10, -6, \frac{1}{4})$. Now, the Hessian of the Lagrangian is

$$H = \begin{pmatrix} 2\lambda & 0 \\ 0 & 2\lambda \end{pmatrix}.$$

So, (x_1, y_1, λ_1) is a minimizer while (x_2, y_2, λ_2) is a maximizer.

Example 5. Consider the following problem:

$$\begin{aligned} \underset{(x_1,x_2)\in\mathbb{R}^2}{\text{minimize}} \quad & f(x_1, x_2) = x_1 + x_2 \\ \text{subject to} \quad & h_1(x_1, x_2) = (x_1 - 1)^2 + x_2^2 - 1 = 0, \\ & h_2(x_1, x_2) = (x_1 - 2)^2 + x_2^2 - 4 = 0 \end{aligned}$$

The solution is $(x_1, x_2) = (0, 0)$. The optimality condition writes:

$$\nabla_x f(x) + \lambda_1 \nabla_x h_1(x) + \lambda_2 \nabla_x h_2(x) = 0.$$

However, the gradients at the origin are $\nabla_x f(0, 0) = (1, 1)$, $\nabla_x h_1(0, 0) = (-2, 0)$, $\nabla_x h_2(0, 0) = (-4, 0)$. Therefore, there are no Lagrange multipliers that enforce the optimality condition.

4 Inequality constrained problem

Consider the following problem:

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & h_i(\boldsymbol{x}) = 0, \forall\, i = 1, \ldots, m, \\ & g_j(\boldsymbol{x}) \leq 0, \forall\, j = 1, \ldots, p. \end{aligned} \tag{II.6}$$

$h_i, g_j : \mathbb{R}^n \to \mathbb{R}$ supposed to be differentiable. We put $\boldsymbol{h} = (h_1, \cdots, h_m)$ and $\boldsymbol{g} = (g_1, \cdots, g_p)$.

When the constraints also have inequalities, the Lagrange Multipliers method is extended to the *KKT conditions*. The expression for the Lagrangian becomes:

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}}{\operatorname{argmin}}\ \mathcal{L}(\boldsymbol{x}, \lambda, \mu) = \underset{\boldsymbol{x}}{\operatorname{argmin}}\ f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i h_i(\boldsymbol{x}) + \sum_{i=1}^{\ell} \mu_i g_i(\boldsymbol{x}), \tag{II.7}$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_\ell)$ are vectors of the multipliers.

Definition 5. A point $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ is called *feasible* when it satisfies all the constraints, i.e. $\boldsymbol{h}(\bar{\boldsymbol{x}}) = 0$ and $\boldsymbol{g}(\bar{\boldsymbol{x}}) \leqslant 0$. We say that a feasible point $\boldsymbol{x}$ is *regular* when $\nabla h_1(\boldsymbol{x})$, $\dots$, $\nabla h_m(\boldsymbol{x})$ are linearly independent.

Definition 6. An inequality constraint $g_j(\boldsymbol{x}) \leqslant 0$ is said to be active at $\boldsymbol{x}^*$ if $g_j(\boldsymbol{x}^*) = 0$ and inactive if $g_j(\boldsymbol{x}^*) < 0$.

> **Remark:** Equality constraint $h_i(\boldsymbol{x}) = 0$ is considered to be always active.

Definition 7. A feasible point $\boldsymbol{x}$ is said to be regular if gradients of all active constraints are linearly independent.

As previously, we give necessary and sufficient conditions for optimality of problem (II.6).

Theorem 8 (Karush-Kuhn-Tucker (KKT) theorem)**.** If $\boldsymbol{x}$ is a regular point and a local minimizer for the problem (II.6), then there exists $\boldsymbol{\lambda} \in \mathbb{R}^m$, $\boldsymbol{\mu} \in \mathbb{R}^p$ such that

- Stationarity: $\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \lambda, \mu) = 0$.
- Dual feasibility: $\boldsymbol{\mu} \geq 0$.
- Complementary slackness: $\mu_j g_j(\boldsymbol{x}) = 0, \forall j$, or, equivently, $\boldsymbol{\mu}^T g(\boldsymbol{x}) = 0$.

> **Remark:** The fact that $\boldsymbol{x}$ is supposed to be a regular point adds the constraints:
>
> $$\boldsymbol{h}(\boldsymbol{x}) = 0; \quad \boldsymbol{g}(\boldsymbol{x}) \leqslant 0.$$

Theorem 9 (Second order conditions)**.** If $\boldsymbol{x}$ is a regular point and a local minimizer for the problem (II.6), then the Hessian of the Lagrangian, $H = \nabla^2_{\boldsymbol{x}\boldsymbol{x}}(\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$, is positive semidefinite on the tangent space:

$$T(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^n \ : \ D\boldsymbol{h}(\boldsymbol{x})\boldsymbol{y} = 0, \ Dg_j(\boldsymbol{x})\boldsymbol{y} = 0 \text{ when } g_j(\boldsymbol{x}) = 0.\}$$

Conversely, if H is positive definite on the space:

$$\bar{T}(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^n \ : \ D\boldsymbol{h}(\boldsymbol{x})\boldsymbol{y} = 0, \ Dg_j(\boldsymbol{x})\boldsymbol{y} = 0 \text{ when } g_j(\boldsymbol{x}) = 0, \mu_j > 0\},$$

and the first order conditions are satisfied, the $\boldsymbol{x}$ is a local minimizer.

Example 6. Consider the following minimization problem

$$\begin{aligned} \underset{\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2}{\text{minimize}} \quad & \|\boldsymbol{x}\|_2 \\ \text{subject to} \quad & x_1 + x_2 + 1 \ \leq 0 \end{aligned}$$

We find the solution of the KKT conditions.

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \sqrt{x_1^2 + x_2^2} + \mu(x_1 + x_2 + 1)$$

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_1} = \frac{x_1}{\|\boldsymbol{x}\|_2} + \mu = 0 \\ \frac{\partial \mathcal{L}}{\partial x_2} = \frac{x_2}{\|\boldsymbol{x}\|_2} + \mu = 0 \\ \mu(x_1 + x_2 + 1) = 0 \\ x_1 + x_2 + 1 \quad \leq 0 \\ \mu \quad \geq 0 \end{cases}$$

So, the unique stationary point is $\boldsymbol{x}^* = -\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$. It is easy to check the second order sufficient condition (first, determine $\bar{T}(\boldsymbol{x})$). x^* is a minimizer.

Example 7. Solve the following minimization problem

$$\begin{aligned} \underset{\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2}{\text{minimize}} \quad & 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 \\ \text{subject to} \quad & x_1^2 + x_2^2 \leqslant 5, \\ & 3x_1 + x_2 \leqslant 6 \end{aligned}$$

The solution is $x_1 = 1$, $x_2 = 2$, $\lambda_1 = 1$, $\lambda_2 = 0$.

> **Remark:** Strict inequalities cannot be taken into account in KKT conditions. To solve the problem, one need to compute stationary points of the problem without these constraints, then throw away all of them that violate these conditions.

5 Least squares

Consider the system of linear equations

$$A\mathbf{x} = \mathbf{y}, \tag{II.8}$$

where $A \in \mathbb{R}^{n \times m}$, $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$, and $n \neq m$, i.e., the problem is ill-posed. We give in the following some techniques to solve this problem.

5.1 Overdetermined linear systems

Assume that $n > m$ (more observations than variables). The problem has no solution if the data are not redundant. In this case, it is common to seek a solution $\mathbf{x}$ minimizing a Least squares loss:

$$J(\mathbf{x}) = \|\mathbf{y} - A\mathbf{x}\|_2^2.$$

Setting the derivative to zero, we obtain the **normal equation**:

$$\frac{\partial}{\partial \mathbf{x}} J(\mathbf{x}) = 0 \implies A^T A\mathbf{x} = A^T\mathbf{y}$$

When $A^T A$ is invertible, then the solution:

$$\mathbf{x} = \left(A^T A\right)^{-1} A^T,$$

$$\mathbf{x} = A^+ \mathbf{y}, \tag{II.9}$$

where A^+ is the pseudinverse. This solution is also valid when $A^T A$ is not invertible (using the SVD decomposition, we can compute A^+).

5.2 Under-constrained linear systems

Assume that $n < m$ (more variables than observations). The problem may have infinite solutions and needs to be constrained.

For instance, we can be interested in finding the one with least ℓ_2-norm, i.e., solution of the following problem:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \|\boldsymbol{x}\|_2^2, \\ \text{subject to} \quad & \boldsymbol{y} - A\boldsymbol{x} = 0. \end{aligned}$$

Setting the problem in Lagrangian form:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \|\boldsymbol{x}\|_2^2 + \lambda^T(\boldsymbol{y} - A\boldsymbol{x}) = \|\boldsymbol{x}\|_2^2 + \lambda^T\boldsymbol{y} - (A^T\lambda)^T\boldsymbol{x}.$$

The optimality conditions yield:

$$\begin{cases} 2\boldsymbol{x} - A^T\boldsymbol{\lambda} = 0 \\ \boldsymbol{y} \qquad\qquad = A\boldsymbol{x} \end{cases}$$

If AA^T is invertible, we get: $\boldsymbol{\lambda} = 2(AA^T)^{-1}\boldsymbol{y}$ and

$$\boldsymbol{x} = A^T(AA^T)^{-1}\boldsymbol{y},$$

i.e.,

$$\boldsymbol{x} = A^+\boldsymbol{y}.$$

Again, this solution is also valid when $A^T A$ is not invertible.

Tikhonov regularization

To cope with the ill-posedness of the problem II.8, another approach is to minimize the weighted sum

$$J(\boldsymbol{x}) = \|\boldsymbol{y} - A\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_2^2,$$

with $\lambda > 0$. Setting the derivative to zero,

$$\boldsymbol{x} = \left(A^T A + \lambda I\right)^{-1} A^T\boldsymbol{y}.$$

This is referred to as "diagonal loading" because a constant, λ, is added to the diagonal elements of $A^T A$. The matrix $A^T A + \lambda I$ is invertible even if $A^T A$ is not. Indeed, it is easy to show that if α is a eigenvalue of a matrix M, the $\alpha + \lambda$ is an eigenvalue of $M + \lambda I$. As $A^T A$ is symmetric, all its eigenvalues are real. So, the eigenvalues of $A^T A + \lambda I$ are positive. Hence, this matrix is invertible.

6 Applications

6.1 Data fitting

We want to fit a line $y = ax + b$ to a given point cloud $(x_1, y_1), \ldots, (x_n, y_n)$ of $\mathbb{R}^2$.

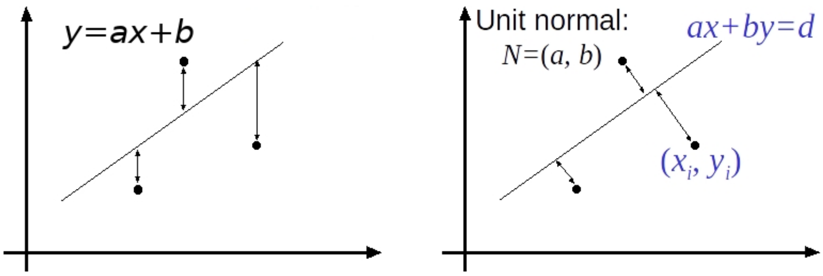

Fig. II.2: Least squares fitting: orthogonal (left) and normal (right).

So, we define a data term

$$E(\boldsymbol{c}) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 = \|A\boldsymbol{c} - \mathrm{d}\|_2^2,$$

where

$$A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} ; \mathbf{c} = \begin{bmatrix} a \\ b \end{bmatrix} \; ; \; \mathbf{d} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} .$$

This model is called *Orthogonal Least Squares*. The minimization of E leads to an optimal estimation of the fitting line. Nevertheless, such a model fails for vertical lines and is not rotation-invariant. To cope with that, we can slightly modify the distance between the points and the estimated line (See Figure II.2), where the line is represented by an equation of the form:

$$ax + by = d.$$

We define the associated data term

$$J(a, b, d) = \sum_{i=1}^{n} (ax_i + by_i - d)^2.$$

This model is called *Total least squares*.

We can eliminate d:

$$\frac{\partial J}{\partial d} = \sum_{i=1}^{n} -2(ax_i + by_i - d) = 0.$$

Put

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \; , \; \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i.$$

We get,

$$\begin{aligned}
E(a,b) &= \sum_{i=1}^{n} (a(x_i - \bar{x}) + b(y_i - \bar{y}))^2 \\
&= \left\| \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \right\|_2^2 \\
&= (U\boldsymbol{c})^T (U\boldsymbol{c}) \\
&= \boldsymbol{c}^T (U^T U) \boldsymbol{c}
\end{aligned}$$

To ensure uniqueness of the solution, we formulate the problem:

$$\begin{aligned}
\underset{\boldsymbol{c}=(a,b)\in\mathbb{R}^2}{\text{minimize}} \quad & E(a,b), \\
\text{subject to} \quad & \|\boldsymbol{c}\|_2^2 = 1.
\end{aligned}$$

We can show that the solution of the latter problem is the eigenvector of $U^T U$ associated with the smallest eigenvalue.

6.2 Signal denoising

Suppose that a noisy measurement of a signal $x \in \mathbb{R}^n$ is given:

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\epsilon}.$$

where $\boldsymbol{x}$ is an unknown signal, $\boldsymbol{\epsilon}$ is an unknown noise vector, and $\boldsymbol{y}$ is the known measurements vector. To denoise the signal $\boldsymbol{b}$ and so find a "good" estimate of $\boldsymbol{x}$, we consider a least squares problem, for wich we add a regularization term R:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{x} - \boldsymbol{y}\|^2 \; + \; \lambda \, R(\boldsymbol{x}), \tag{II.10}$$

where λ is a given regularization para- meter and $R(\boldsymbol{x})$ is sum of the squares of the differences of consecutive components of $\boldsymbol{x}$:

$$R(\boldsymbol{x}) = \sum_{i=1}^{n-1} (x_i - x_{i+1})^2.$$

This quadratic function can also be written as $R(\boldsymbol{x}) = \|\boldsymbol{L}\boldsymbol{x}\|^2$, where $\boldsymbol{L} \in \mathbb{R}^{(n-1)\times n}$ is given by

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

The optimal solution of (II.10) is given by

$$\bar{\boldsymbol{x}} = (\boldsymbol{I} + +\lambda \boldsymbol{L}^T \boldsymbol{L})^{-1} \boldsymbol{y}.$$

6.3 Image denoising

Consider a noisy image $Y \in \mathbb{R}^{n\times n}$ and its associated vectorial represntation $\boldsymbol{y}$. As in the previous section, we express the denoising of $\boldsymbol{y}$ as the following least squares problem:

$$\min_{\boldsymbol{x}\in\mathbb{R}^N} \|\boldsymbol{x}-\boldsymbol{y}\|^2 \ + \ \lambda\, R(\boldsymbol{x}), \tag{II.11}$$

where λ is a given regularization para- meter, $N = n^2$, and $R(\boldsymbol{x})$ is a regularization term

$$R(\boldsymbol{x}) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n-1}(x_{i,j}-x_{i+1,j})^2 + (x_{n,j}-x_{i-1,j})^2\right).$$

The regularization term R corresponds to the $x-$derivative only. It can also be written as $R(\boldsymbol{x}) = \|\boldsymbol{L}\boldsymbol{x}\|^2$, where $\boldsymbol{L} \in \mathbb{R}^{N\times N}$ is given by

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{I} & 0 & \cdots & 0 \\ 0 & \boldsymbol{I} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{I} \end{bmatrix}, \quad \text{where} \quad \boldsymbol{I} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}$$

Again, the optimal solution of (II.13) is given by

$$\bar{\boldsymbol{x}} = (\boldsymbol{I} + \lambda \boldsymbol{L}^T\boldsymbol{L})^{-1}\boldsymbol{y}.$$