Technical University of Denmark

Computational Imaging and Spectroscopy: Linear inverse problems

Thierry SOREZE

DTU July 2021

DTU Fotonik

Department of Photonics Engineering

Linear inverse problems

Problem formulation

Many problems in image processing can be formulated as the inversion of the system:

$$y = \mathbf{H}x_0 \odot \varepsilon$$

Where $x_0 \in \mathbb{R}^N$, is the signal we seek to recover, $y \in \mathbb{R}^m$ is the vector of corrupted observations. $\odot$ is either $+$ or $\times$, and ε is an unknown noise model. $\mathbf{H} : \mathbb{R}^N \rightarrow \mathbb{R}^m$ is a linear operator, typically ill behaved.

This problem is generally ill-posed, and need to be regularized, by reducing the space of candidate solutions, by adding prior knowledge on the structure of the unknown vector x_0

Linear inverse problems

Problem formulation

Deconvolution: $\mathbf{H}$ is the convolution by a blurring kernel, y lacks the high frequency components of x_0

Inpainting: $\mathbf{H}$ is a pixelwise multiplication by a binary mask

Decoding (Compressed Sensing): $\mathbf{H}$ is an $m \times N$ sensing matrix taking $m \ll N$ measurements at random from the input signal x_0, supposed to be sparse in a dictionary $\mathbf{\Phi}$

Linear inverse problems

Sparsity regularized inverse problems

Here we assume that the solution of our initial problem is sparsely represented in some dictionary $\mathbf{\Phi}$

The problem we which to solve is then cast as the following composite and structured minimization problem (P):

$$\min_{x \in \mathbb{R}^N} D(\mathbf{H}x, y) + \sum_{k=1}^{l} R_k(x)$$

$D: \mathbb{R}^m \times \mathbb{R}^m$ is a function measuring the consistency to the observed data y , and R_k are functions encoding the priors to be imposed on the signal to be recovered, x_0

We consider $D(\cdot, y), \forall y,$ and $R_k \forall k$ to be lower semi continuous convex functions. R_1 is usually a sparsity promotion constraint

Linear inverse problems

Sparsity regularized inverse problems

i. Synthesis Sparsity problems

We seek a sparse set of coefficients α and solution image which is synthesized from these coefficients as $x = \mathbf{\Phi}\alpha$, where $\mathbf{\Phi} \in \mathbb{R}^{N \times T}$. This type of prior is called synthesis sparsity prior.

The ℓ_1 decoder known as **Basis Pursuit** in the literature reads:

$$\min_{\alpha \in \mathbb{R}^T} \|\alpha\|_1 \qquad \text{s.t} \quad y = \mathbf{H\Phi}\alpha = \mathbf{F}\alpha$$

This is an instance of (P) in which $l = 1$

$$R_1(x) = \min_{\alpha \in \mathbb{R}^T} \|\alpha\|_1 \qquad \text{s.t} \quad x = \mathbf{\Phi}\alpha$$

Linear inverse problems

Sparsity regularized inverse problems

i. Synthesis Sparsity problems

$D(\mathbf{H}\,\cdot, y)$ is the indicator function of the affine subspace $\{x \in \mathbb{R}^N | y = \mathbf{H}x\}$

In presence of noise, the equality constraint must be relaxed to a noise aware variant. (P) becomes Basis Pursuit DeNoising (BPDN) when

$$\min_{\alpha \in \mathbb{R}^T} \frac{1}{2} \|y - \mathbf{F}\alpha\|^2 + \lambda \|\alpha\|_1, \qquad \lambda > 0$$

This setting is also known as Lasso in the literature, in the ℓ_1 constrained form

$$\min_{\alpha \in \mathbb{R}^T} \frac{1}{2} \|y - \mathbf{F}\alpha\|^2 \ \ \mathrm{s.t} \qquad \|\alpha\|_1 \leq \rho$$

Linear inverse problems

Sparsity regularized inverse problems

i. Synthesis Sparsity problems

$D(\mathbf{H}x, y) = \frac{1}{2}\|y - \mathbf{H}x\|^2$, and R_1 is up to a multiplication by λ

These formulations are equivalent to the constrained form

$$\min_{\alpha \in \mathbb{R}^T} \|\alpha\|_1 \qquad \mathrm{s.t} \quad \|y - \mathbf{F}\alpha\| \leq \sigma$$

$D(\mathbf{H}\,\cdot, y)$ is the indicator function of the closed convex space $\{x \in \mathbb{R}^N |\, \|y =$

Linear inverse problems

Sparsity regularized inverse problems

i. Synthesis Sparsity problems

The Danzig selector corresponds to:

$$\min_{\alpha \in \mathbb{R}^T} \|\alpha\|_1 \qquad \text{s.t} \quad \left\|\mathbf{F}^{\mathrm{T}}(y - \mathbf{F}\alpha)\right\| \leq \delta$$

$D(\mathbf{H}\,\cdot, y)$ is the indicator function of $\left\{x \in \mathbb{R}^N | \left\|\mathbf{F}^{\mathrm{T}}(y - \mathbf{H}x)\right\|_\infty \leq \delta\right\}$

Linear inverse problems

Sparsity regularized inverse problems

ii. Analysis Sparsity problems

In the analysis sparsity prior framework, we seek a solution image x whose coefficients $\mathbf{\Phi}^{\mathrm{T}}x$ are sparse. The ℓ_1-analysis prior formulation of the previous equations are given by:

$$\min_{x\in\mathbb{R}^N} \left\|\mathbf{\Phi}^{\mathrm{T}}x\right\|_1 \qquad \mathrm{s.t} \quad y = \mathbf{H}x$$

$$\min_{x\in\mathbb{R}^N} \frac{1}{2}\|y-\mathbf{H}x\|^2 + \lambda\left\|\mathbf{\Phi}^{\mathrm{T}}x\right\|_1$$

Linear inverse problems

Sparsity regularized inverse problems

ii. Analysis Sparsity problems

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|y - \mathbf{H}x\|^2 + \quad \mathrm{s.t} \quad \left\|\boldsymbol{\Phi}^{\mathrm{T}} x\right\|_1 \leq \rho$$

$$\min_{x \in \mathbb{R}^N} \left\|\boldsymbol{\Phi}^{\mathrm{T}} x\right\|_1 \quad \mathrm{s.t} \quad \|y - \mathbf{H}x\| \leq \sigma$$

$$\min_{x \in \mathbb{R}^N} \left\|\boldsymbol{\Phi}^{\mathrm{T}} x\right\|_1 \quad \mathrm{s.t} \quad \left\|\mathbf{H}^{\mathrm{T}}(y - \mathbf{H}x)\right\|_\infty \leq \delta$$

Linear inverse problems

Basics of Convex Analysis

Let $\mathcal{H}$ be a finite dimensional real vector space, equipped with the inner product $\langle \cdot\ ,\ \cdot \rangle$ and associated norm $\|\cdot\|$

Let $\mathbf{I}$ be the identity operator on $\mathcal{H}$

- The operator spectral norm of the linear operator $\mathbf{A}: \mathcal{H} \to \mathcal{K}$, with $\mathcal{K}$ a finite dimensional real vector space, is denoted $|||\mathbf{A}||| = \sup_{x \in \mathcal{H}} \frac{\|\mathbf{A}\|}{\|x\|}$

- $\|x\|_p = (\sum_i |x[i]|^p)^{1/p}$ is the ℓ_p norm, $\|x\|_\infty = \max_i |x[i]|$

- $\mathbb{B}_p^\rho$ is the closed ℓ_p ball of radius $\rho > 0$ centered at the origin

Linear inverse problems

Basics of Convex Analysis

- A real valued function $F : \mathcal{H} \to (-\infty, +\infty]$ is coercive if $\lim_{\|x\| \to +\infty} F(x) = +\infty$

- The domain of F is defined by $\mathrm{dom}\, F = \{x \in \mathcal{H} \mid F(x) < +\infty\}$

- F is proper if $\mathrm{dom}\, F \neq \emptyset$

- F is semi lower-continuous (lsc) if $\lim_{\|x\| \to 0} F(x) \geq F(x_0)$

Nota: Semi lower-continuity is weaker than continuity

- $\Gamma_0(\mathcal{H})$ is the class of all proper lsc convex functions $\mathcal{H} \to (-\infty, +\infty]$

Linear inverse problems

Basics of Convex Analysis

Let $\mathcal{C}$ be an non-empty convex subset of $\mathcal{H}$

- $\mathrm{ri}(\mathcal{C})$ denotes its relative interior

- The indicator function $I_{\mathcal{C}}$ of $\mathcal{C}$ is

$$I_{\mathcal{C}}(x) = \begin{cases} 0, & \text{if } x \in \mathcal{C} \\ +\infty, & \text{othertwise} \end{cases}$$

- The conjugate of as function $F \in \Gamma_0(\mathcal{H})$ is F^*, defined by

$$F^*(u) = \sup_{x \in \mathrm{dom}\, F} \langle u, x \rangle - F(x)$$

Linear inverse problems

Basics of Convex Analysis

The subdifferential of a function $F \in \Gamma_0(\mathcal{H})$ at $x \in \mathcal{H}$ is the set-valued map ∂F from $\mathcal{H}$ into subsets of $\mathcal{H}$:

$$\partial F(x) = \{u \in \mathcal{H} \mid \forall z \in \mathcal{H}, F(z) \geq F(x) + \langle u, z - x \rangle\}$$

An element of $\partial F(x)$ is called a subgradient. If F is differentiable at x, its only subgradient is its gradient, i.e. $\partial F(x) = \nabla F(x)$

An everywhere differentiable function has a Lipschitz β-continuous gradient, $\beta \geq 0$, if

$$\|\nabla F(x) - \nabla F(z)\| \leq \beta \|x - z\|, \qquad \forall (x, z) \in \mathcal{H}^2$$

Linear inverse problems

Basics of Convex Analysis

A function is strictly convex if the latter inequality holds as a strict inequality for $z \neq x$

A function is strongly convex with modulus $c > 0$ if and only if

$$F(z) \geq F(x) + \langle u, z + x \rangle + \frac{c}{2} \|z - x\|^2, \forall z \in \mathcal{H}$$

x^* is a global minimizer of $F \in \Gamma_0\,(\mathcal{H})$ over $\mathcal{H}$, if and only if

$$0 \in \partial F(x^*)$$

This minimum is unique if F is strictly convex

Linear inverse problems

Proximal calculus

Proximity operator

Let $F \in \Gamma_0\ (\mathcal{H})$

For every $\alpha \in \mathcal{H}$, the function $z \to \frac{1}{2}\|\alpha - z\|^2 + F(x)$ achieves its infimum at a unique point denoted by $\mathrm{prox}_F(\alpha)$.

The uniquely valued operator $\mathrm{prox}_F : \mathcal{H} \to \mathcal{H}$ is the proximity operator of F

When $F = I_{\mathcal{C}}$ for a closed convex set $\mathcal{C}$, prox_F is the projector onto $\mathcal{C}$

An example of proximity operator, is the one associated with $F(\alpha) = \lambda\|\alpha\|_1$ for $\alpha \in \mathbb{R}^{\mathrm{T}}$ is the soft-thresholding with threshold λ

Linear inverse problems

Proximal calculus

Proximity operator

Let $F \in \Gamma_0\ (\mathcal{H})$

For every $\alpha \in \mathcal{H}$, the function $z \to \frac{1}{2}\|\alpha - z\|^2 + F(x)$ achieves its infimum at a unique point denoted by $\mathrm{prox}_F(\alpha)$.

The uniquely valued operator $\mathrm{prox}_F : \mathcal{H} \to \mathcal{H}$ is the proximity operator of F

When $F = I_{\mathcal{C}}$ for a closed convex set $\mathcal{C}$, prox_F is the projector onto $\mathcal{C}$

An example of proximity operator, is the one associated with $F(\alpha) = \lambda\|\alpha\|_1$ for $\alpha \in \mathbb{R}^{\mathrm{T}}$ is the soft-thresholding with threshold λ

Linear inverse problems

Proximal calculus

Proximity operator

i. $\forall z \in \mathcal{H}$, $\mathrm{prox}_{F(\cdot - z)}(\alpha) = z + \mathrm{prox}_F(\alpha - z)$

ii. $\forall z \in \mathcal{H}$, $\forall \rho \in (-\infty, \infty)$, $\mathrm{prox}_{F(\rho\cdot)}(\alpha) = \mathrm{prox}_{\rho^2 F}(\rho\alpha)/\rho$

iii. $\forall z \in \mathcal{H}$, $\forall \rho > 0$, $\tau \in \mathbb{R}$, let $G(\alpha) = F(\alpha) + \rho\|\alpha\|^2 + \langle \alpha, z\rangle + \tau$ Then $\mathrm{prox}_G = \mathrm{prox}_{F/(1+\rho)((\alpha - z)/(\rho+1))}$

iv. *Separability:* let $\{F_i\}_{1 \le i \le n}$ be a family of functions in $\Gamma_0(\mathcal{H})$ and F defined on $\mathcal{H}^n$ with $F\{\alpha_1, \cdots, \alpha_n\} = \sum_{i=1}^n F_i(\alpha_i)$. Then $\mathrm{prox}_F(\alpha) = \left(\mathrm{prox}_{F_i}(\alpha)\right)_{1 \le i \le n}$

We also have

$$\mathrm{prox}_{F^*} = \mathbf{I} - \mathrm{prox}_F$$

Linear inverse problems

Proximal calculus

Proximity operator of a convex Sparsity Penalties

Here we consider a family of simple penalties , i.e. their proximity operator has a simple closed form:

$$J(\alpha) = \sum_{i=1}^{T} \psi_i(\alpha[i])$$

We assume $\forall 1 \leq i \leq T$*:*

Linear inverse problems

Proximal calculus

Proximity operator of a convex Sparsity Penalties

i. $\psi_i \in \Gamma_0\ (\mathcal{H})$

ii. ψ_i *is even symmetric, non negative, non decreasing on* $[0, +\infty)$

iii. ψ_i *is continuous on* $\mathbb{R}$, *with* $\psi_i(0) = 0$

iv. ψ_i *is differentiable on* $(0, +\infty)$, *but non necessarily smooth at 0 and admits a positive right derivative at zero*

We have that (c.f property iv)

$$\mathrm{prox}_{\lambda J}(\alpha) = \left(\mathrm{prox}_{\lambda\psi_i}(\alpha[i])\right)_{1 \le i \le T}, \qquad \lambda > 0$$

Linear inverse problems

Proximal calculus

Proximity operator of a convex Sparsity Penalties

Where $\tilde{\alpha}[i] \equiv \mathrm{prox}_{\lambda\psi_i}(\alpha[i])$, has exactly one continuous and odd-symmetric solution:

$$\tilde{\alpha}[i] = \begin{cases} 0 & \text{if } |\alpha[i]| \leq \lambda\psi_{i+}^{\prime\prime}(0) \\ \alpha[i] - \lambda\psi_i^{\prime\prime}(\tilde{\alpha}[i]) & \text{if } |\alpha[i]| > \lambda\psi_{i+}^{\prime\prime}(0) \end{cases}$$

Linear inverse problems

Proximal calculus

Proximity operator of a convex Sparsity Penalties

Example:

$$\psi_i : t \in \mathbb{R} \longmapsto \lambda |t|^p \,, p \geq 1$$

For p=1 this corresponds to $J(\alpha) = \lambda \|\alpha\|_1,$ whose proximity operator is given by soft-thresholding

$$\mathrm{prox}_{\lambda\|\cdot\|_1 = \mathbf{SoftThreshold}_\lambda(\alpha)} = \left(\left(1 - \frac{\lambda}{|\alpha[i]|} \right)_+ \alpha[i] \right)_{1 \leq i \leq T}$$

$$(\cdot)_+ = \max(\cdot, 0)$$

Linear inverse problems

Proximal calculus

Projection on the ℓ_p ball

For $\rho > 0$ *and* $p \geq 1$, $\mathbb{B}_p^\rho$ is the closed ℓ_p ball of radius ρ

- $p = 2$

$$\mathbf{P}_{\mathbb{B}_2^\rho}(\alpha) = \begin{cases} \alpha & if \;\; \|\alpha\| \leq \rho \\ \alpha\rho/\|\alpha\| & \text{otherwise} \end{cases}$$

Linear inverse problems

Proximal calculus

Projection on the ℓ_p ball

For $\rho > 0$ *and* $p \geq 1$, $\mathbb{B}_p^\rho$ is the closed ℓ_p ball of radius ρ

- $p = \infty$

$$\mathbf{P}_{\mathbb{B}_\infty^\rho}(\alpha) = \left(\frac{\alpha[i]}{\max\left(\frac{|(\alpha[i])|}{\rho}, 1\right)} \right)_{1 \leq i \leq T}$$

Linear inverse problems

Proximal calculus

Projection on the ℓ_p ball

For $\rho > 0$ *and* $p \geq 1$, $\mathbb{B}_p^\rho$ is the closed ℓ_p ball of radius ρ

- $p = 1$, if $\|\alpha\| \leq \rho$, then $\mathbf{P}_{\mathbb{B}_1^\rho}(\alpha) = \alpha$, otherwise $\mathbf{P}_{\mathbb{B}_1^\rho}(\alpha)$ can be computed with $\mathbf{SoftThreshold}_\lambda(\alpha)$

The Langrage multiplier $\lambda(\rho)$ can be obtained by

$$\lambda(\rho) = \alpha_{(j)} + \left(\alpha_{(j+1)} - \alpha_{(j)}\right)\frac{\tilde{\alpha}_{j+1} - \rho}{\tilde{\alpha}_{j+1} + \tilde{\alpha}_j}$$

Linear inverse problems

Proximal calculus

Affine Subspace Projector

Let $\mathbf{F}$ be a linear operator. Let y be in the range of $\mathbf{F}$. We seek to compute the projector on the affine subspace $\mathcal{C} = \{\alpha \in \mathbb{R}^T \,|\mathbf{F}\alpha = y\}$. We have

$$\mathbf{P}_{\mathcal{C}}(\alpha) = \alpha + \mathbf{F}^{+}(y - \mathbf{F}\alpha)$$

Linear inverse problems

Proximal calculus

Pre-composition with an affine Operator

Let $\mathbf{F}$ be a linear operator. Proximity operator of the pre-composition of $F \in \Gamma_0\,(\mathcal{H})$ with the affine mapping $\mathbf{A} : \mathcal{H} \to \mathcal{K}, \alpha \longmapsto \mathbf{F}\alpha$, i.e. solving

$$\min_{z \in \mathcal{H}} \frac{1}{2} \|\alpha - z\|^2 + F(\mathbf{A}z)$$

If $\mathbf{F}$ is orthogonal then

$$\mathrm{prox}_{F \circ \mathbf{A}}(\alpha) = \mathbf{F}^{\mathrm{T}}\big(y + \mathrm{prox}_F(\mathbf{F}\alpha - y)\big)$$

Linear inverse problems

Proximal calculus

Pre-composition with an affine Operator

If $\mathbf{F}$ is a tight frame with constant c, then $F \circ \mathbf{A} \in \Gamma_0\ (\mathcal{H})$ and

$$\mathrm{prox}_{F\circ\mathbf{A}}(\alpha) = \alpha + \ c^{-1}\ \mathbf{F}^{\mathrm{T}}(y + \mathrm{prox}_{cF} - \mathbf{I})(\mathbf{F}\alpha - y)$$

$\mathbf{F}$ *is a general frame, with lower and upper bounds* c_1 *and* $c_2, F \circ \mathbf{A} \in \Gamma_0\ (\mathcal{H})$

Linear inverse problems

Proximal calculus

Algorithm: *Iterative scheme to compute the proximity operator of precomposition with an affine operator*

Init: choose some $u^0 \in \mathrm{dom}(F^*)$, set $\mathrm{p}^0 = \alpha - \mathbf{F}^{\mathrm{T}} u^{(0)}, \mu \in \left]0, 2/c_2\right[$

For t=0 to Niter -1 **do**

$$u^{(t+1)} = \mu\left(\mathbf{I} - \mathrm{prox}_{\mu^{-1}F}\right)\left(\mu^{-1} u^{(t)} + \mathbf{A}\mathrm{p}^t\right)$$

$$\mathrm{p}^{(t+1)} = \alpha - \mathbf{F}^{\mathrm{T}} u^{(t+1)}$$

Linear inverse problems

Proximal Splitting framework

We recall that we seek to minimize functions in the form $F = D(\mathbf{H}\,\cdot, y) + \sum_k R_k$

We assume the set minimizer of (P) to be nonempty. x^* is a global minimizer of (P) if and only if

$$\begin{gathered} 0 \in \partial F(x^*) \\ \Leftrightarrow 0 \in \partial(\gamma F)(x^*), \forall \gamma > 0 \\ \Leftrightarrow x^* - x^* \in \partial(\gamma F)(x^*) \\ \Leftrightarrow x^* = \mathrm{prox}_{\gamma F}(x^*) \end{gathered}$$

Linear inverse problems

Proximal Splitting framework

$\gamma > 0$ is the proximal step size. The proximal type algorithm is constructed as

$$x^{t+1} = \mathrm{prox}_{\gamma F}\left(x^{(t)}\right)$$

The synthesis sparsity prior reads for (P)

$$\min_{\alpha \in \mathbb{R}^T} D(\mathbf{F}\alpha, y) + \sum_{k=1}^{l} J_k(\alpha)$$

Where the regularizing penalties $J_k \in \Gamma_0\ (\mathcal{H})$. For $l = 1$,

$$R_1(x) = \min_{\alpha} J_1(\alpha) \qquad \mathrm{s.t} \qquad x = \mathbf{\Phi}\alpha$$

Linear inverse problems

Proximal Splitting framework

Forward-Backward

$$\begin{cases} \mu \in \left]0, \dfrac{2}{\beta |||\mathbf{F}|||^2}\right[, \tau_t \in [0, \kappa], \kappa = \dfrac{4 - \mu\beta |||\mathbf{F}|||^2}{2} \in]1,2[, \sum_{\boldsymbol{t} \in \mathbb{N}} \tau_t(\kappa - \tau_t) = +\infty \\ \alpha^{(\boldsymbol{t}+\mathbf{1})} = \alpha^{(\boldsymbol{t})} + \tau_t \left(\mathrm{prox}_{\mu J_1}\left(\alpha^{(t)} - \mu \mathbf{F}^{\mathrm{T}} \nabla D\left(\mathbf{F}\alpha^{(\boldsymbol{t})}, y\right)\right) - \alpha^{(t)}\right) \end{cases}$$

Linear inverse problems

ISTA for BPDN

Algorithm: *Iterative Soft-Thresholding Algorithm*

Init: choose some $\alpha^{(0)}$, $\mu \in \left]\frac{2}{|||\mathbf{H}|||^2|||\Phi|||^2}\right[, \tau_t \in [0, \kappa], \kappa = \frac{4-\mu\beta|||\mathbf{H}|||^2|||\Phi|||^2}{2} \in$ $]1,2[, \sum_{t\in\mathbb{N}} \tau_t(\kappa - \tau_t) = +\infty$

For t=0 to Niter -1 **do**

1. Gradient descent: $\alpha^{(t+1/2)} + \mu\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{H}^{T}\left(y - \mathbf{H}\boldsymbol{\Phi}\alpha^{(t)}\right)$
2. Soft-Thresholding: $\alpha^{(t+1)} = \alpha^{(t)} + \tau_t\left(\mathbf{SoftThresh}_{\mu\lambda}\left(\alpha^{(t+1/2)}\right) - \alpha^{(t)}\right)$

Linear inverse problems

Proximal Splitting framework

Splitting scheme	Objective	Asssumption
Forward-backward	$D(\mathbf{F}\cdot, y) + J_1$	J_1 simple, $\nabla D(\mathbf{F}\cdot, y)$ Lipschitz
FISTA	$D(\mathbf{F}\cdot, y) + J_1$	J_1 simple, $\nabla D(\mathbf{F}\cdot, y)$ Lipschitz
Douglas-Rachford	$D(\mathbf{F}\cdot, y) + J_1$	J_1 and $D(\mathbf{F}\cdot, y)$ simple
ADMM	$D(\mathbf{F}\cdot, y) + J_1$	J_1 and $D(\mathbf{F}\cdot, y)$ simple, $\mathbf{F}\mathbf{F}^{\mathrm{T}}$ invertible
GFB	$D(\mathbf{F}\cdot, y) + \sum_{k=1}^{l} J_k$	J_k simple, $\nabla D(\mathbf{F}\cdot, y)$ Lipschitz
Primal-dual	$D(\mathbf{F}\cdot, y) + J_1 + \sum_{k=2}^{l} G_k \circ \mathbf{A}_k$	J_1 and G_k simple, $\mathbf{A}_k$ linear, $\nabla D(\mathbf{F}\cdot, y)$ Lipschitz

Linear inverse problems

Choice of Regularization parameters in BPDN/Lasso

The choice of the parameter λ is of crucial importance in regularized linear inverse problems, as it represents the balance between sparsity and data fidelity.

λ should always be chosen in $\left]0, \|\mathbf{F}^{\mathrm{T}}y\|_{\infty}\right[$. In this setting α^* is a global minimizer of the synthesis sparsity BPND problem, if and only if

$$\mathbf{F}^{\mathrm{T}}(y - \mathbf{F}\alpha^*) \in \lambda\partial\|\cdot\|_1(\alpha^*) \subseteq \lambda\partial\|\cdot\|_1(0) = \mathbb{B}_{\infty}^{\lambda}$$

Assume that the noise corrupting our observations is Gaussian with variance σ_ε^2, the value of λ can be taken as:

$$\lambda = c\sigma_\varepsilon\sqrt{2\log T}$$

If c is large enough, i.e. $c > 2\sqrt{2}$, then BPDN is able to recover the correct sparsity or be consistent with the original sparse vector, under additional conditions on $\mathbf{F}$

Linear inverse problems

Choice of Regularization parameters in BPDN/Lasso

When the noise ε is assumed to be zero mean white Gaussian of variance σ_ε, the SURE provides a framework to select the regularization parameter by minimizing and objective quality measure.

Denote the mapping $\mu_\lambda^* : y \longmapsto \mathbf{F}\alpha_\lambda^*(y)$, where $\alpha_\lambda^*(y)$ is any global minimizer of the BPDN equation with observations y.

We would like to choose λ that minimizes the quadratic risk, for a matrix $\mathbf{A} \in \mathbb{R}^{n\times m}$:

$$\mathbb{E}_\varepsilon\left(\left\|\mathbf{A}_{\mu_0} - \mathbf{A}\alpha_\lambda^*(y)\right\|^2\right)$$

Linear inverse problems

Choice of Regularization parameters in BPDN/Lasso

There are several choices for the matrix $\mathbf{A}$:

$\mathbf{A} = \mathbf{I}$ prediction risk

$\mathbf{F}$ rank deficient, $\mathbf{A} = \mathbf{F}^{\mathrm{T}}\left(\mathbf{F}\mathbf{F}^{\mathrm{T}}\right)^{+}$, projection risk

$\mathbf{F}$ full rank, $\mathbf{A} = \mathbf{F}^{+} = \left(\mathbf{F}^{\mathrm{T}}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathrm{T}}$, estimation risk

In inverse problems as $\mathbf{H}$ is usually singular, the projection risk is preferable

Linear inverse problems

Choice of Regularization parameters in BPDN/Lasso

The Generalized SURE (GSURE) associated with $\mathbf{A}$ is

$$\mathrm{GSURE}^{\mathbf{A}}\left(\mu_\lambda^*(y)\right) = \left\|\mathbf{A}\left(y - \mu_\lambda^*(y)\right)\right\|^2 - \sigma^2 \mathrm{trace}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right) + 2\sigma^2 df^{\mathbf{A}}(y)$$

$$df^{\mathbf{A}}(y) = \mathrm{trace}\left(\mathbf{A}\frac{\partial \mu_\lambda^*(y)}{\partial y}\mathbf{A}^{\mathrm{T}}\right)$$

GSURE is an unbiased estimator of the risk

Linear inverse problems

Application : Linear Solver

Install sparse solvers toolboxes

Run examples with SparCo
Explore the toolboxes and the solvers

Linear inverse problems

Application 1 : denoising

Load Einstein image

Apply Gaussian noise

Cast the denoising problem as a linear inverse problem

Select

i. **A dictionary**
ii. **A solver**

Solve the problem. Study convergence and performance

Linear inverse problems

Application 2 : Inpainting

Load Barbara image

Apply a binary masking

Cast the inpainting problem as a linear inverse problem

Select

i. **A dictionary**
ii. **A solver**

Solve the problem. Study convergence and performance