Iterative methods for Image Processing

Lothar Reichel

Como, May 2018.

Lecture 4: Image restoration based on nonconvex optimization.

Outline of Lecture 4:

- ℓ_p-ℓ_q minimization methods
- Choice of solution subspace
- Selection of regularization parameter

The minimization problem

$$\min_{x\in\mathbf{R}^n} \mathcal{J}(x), \qquad \mathcal{J}(x) = \frac{1}{p}\|Ax - b\|_p^p + \frac{\mu}{q}\|\Phi(x)\|_q^q \,,$$

where

$$0 < p, q \leq 2, \quad \mu > 0,$$
$$A \in \mathbf{R}^{m\times n}, \quad b \in \mathbf{R}^m, \quad x \in \mathbf{R}^n, \quad \Phi : \mathbf{R}^n \to \mathbf{R}^s.$$

Special case: Tikhonov regularization

$$p = 2, \quad q = 2, \quad \Phi(x) = Lx, \;\; L \in \mathbf{R}^{s\times n}.$$

Large-scale problems often solved by Krylov subspace methods.

Applications:

- $p = 2$, $m < n$, $0 < q \leq 1$, and $\Phi = I$: Compute sparse solutions of undetermined linear systems.
- $p = 2$, $0 < q \leq 1$, $\Phi = I$, and A a sampling operator: compressed sensing,
- Image restoration: Each element of x represents a pixel.

ℓ^q-norms: solid black graph: ℓ^0-norm; dotted black graph: ℓ^1-norm; dark gray solid graph: $\ell^{0.5}$-norm; light gray solid graph: $\ell^{0.1}$-norm.

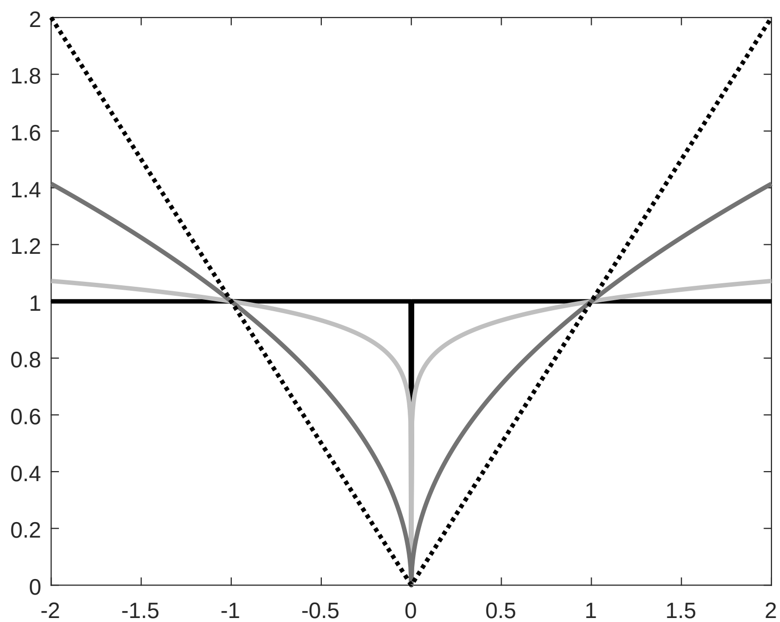


Image restoration applications:

- Total variation (TV) regularization when $q = 1$ and $\Phi = \Phi_{TV} : \mathbf{R}^n \to \mathbf{R}^n$:

$$\|x\|_{\mathrm{TV}} := \|\Phi_{\mathrm{TV}}(x)\|_1 = \sum_{i=1}^{n} \|\left(\nabla x\right)_i\|_2 \ ,$$

where $\left(\nabla x\right)_i \in \mathbf{R}^2$ is the discrete gradient at pixel i.

Gives ℓ_p-TV restoration:

$$\min_{x \in \mathbf{R}^n} \left\{ \frac{1}{p}\|Ax - b\|_p^p + \mu\|x\|_{\mathrm{TV}} \right\} .$$

Use $p = 2$ for additive Gaussian noise, $0 < p \leq 1$ for impulse noise.

Majorization-minimization (MM) methods

Let $G(x) : \mathcal{R}^n \to \mathcal{R}$ be continuously differentiable. The function $Q(x, v) : \mathcal{R}^n \times \mathcal{R}^n \to \mathcal{R}$ is said to be a quadratic tangent majorant for $G(x)$ iff for any $v \in \mathcal{R}^n$ the following conditions hold:

1. $Q(x, v)$ is quadratic in x,
2. $Q(v, v) = G(v)$,
3. $\nabla_x Q(v, v) = \nabla_x G(v)$,
4. $Q(x, v) \geq G(v) \qquad \forall \;\; x \in \mathcal{R}^n$.

Quadratic majorization possible only for continuously differentiable functions. We therefore use smoothed functions

$$\phi_{z,\varepsilon}(t) := \left(\sqrt{t^2+\varepsilon^2}\right)^z \quad \text{with} \quad \begin{cases} \varepsilon > 0 & \text{for} \quad 0 < z \leq 1, \\ \varepsilon = 0 & \text{for} \quad 1 < z \leq 2. \end{cases}$$

Then

$$\phi'_{z,\varepsilon}(t) := \frac{d}{dt}\phi_{z,\varepsilon}(t) = z\,t\left(\sqrt{t^2+\varepsilon^2}\right)^{z-2} = \ z\,t\ \phi_{z-2,\varepsilon}(t)\ .$$

We consider the minimization problem

$$\min_{x \in \mathbf{R}^n} \mathcal{J}_\varepsilon(x),$$

where

$$\mathcal{J}_\varepsilon(x) := \frac{1}{p} \sum_{i=1}^{r} \phi_{p,\varepsilon}\big((Ax-b)_i\big) + \frac{\mu}{q} \sum_{j=1}^{s} \phi_{q,\varepsilon}\big((Lx)_j\big) .$$

Proposition: Let $\phi_{z,\varepsilon}(t) : \mathbf{R} \to \mathbf{R}_+$ be the smoothed penalty function defined above with $z \in]\,0, 2\,]$. Then any function $m_{z,\varepsilon}(t, v) : \mathbf{R} \times \mathbf{R} \to \mathbf{R}_+$ of the form

$$m_{z,\varepsilon}(t, v) \;:=\; a_v\,(t - b_v)^2 + c_v$$

with

$$a_v \;\in\; [\,\underline{a_v}\,, +\infty[\,, \qquad \underline{a_v} \;:=\; \frac{\phi'_{z,\varepsilon}(v)}{2v} \;=\; \frac{z}{2}\;\phi_{z-2,\varepsilon}(v)\;,$$

$$b_v \;:=\; v - \frac{\phi'_{z,\varepsilon}(v)}{2a_v} \;=\; v\left(1 - \underline{a_v}\,/\,a_v\right)\;,$$

$$c_v \;:=\; \phi_{z,\varepsilon}(v) - \frac{\left(\phi'_{z,\varepsilon}(v)\right)^2}{4a_v} \;=\; \phi_{z,\varepsilon}(v) - v^2\,\underline{a_v}^{\,2}\,/\,a_v\;,$$

is a quadratic tangent majorant for $\phi_{z,\varepsilon}(t)$.

Thus,

$$
\begin{aligned}
m_{z,\varepsilon}(v,v) &= \phi_{z,\varepsilon}(v) \qquad \forall\, v \in \mathbf{R}\,, \\
m'_{z,\varepsilon}(v,v) &= \phi'_{z,\varepsilon}(v) \qquad \forall\, v \in \mathbf{R}\,, \\
m_{z,\varepsilon}(t,v) &\geq \phi_{z,\varepsilon}(t) \qquad \forall\, v \in \mathbf{R}\,, \quad \forall\, t \in \mathbf{R}\,.
\end{aligned}
$$

Adaptive and fixed quadratic majorants

Let $\phi_{z,\varepsilon}(t)$ be the smoothed penalty function defined above with $z \in \,]\,0, 2\,]$ and let $m_{z,\varepsilon}(t, v)$ be the family of associated quadratic tangent majorants.

- If for every $v \in \mathbf{R}$, the parameter a_v is chosen as the lower limit of the admissible interval, i.e., if $a_v := \underline{a_v}$, then the quadratic majorant takes the adaptive form:

$$\begin{aligned} m^{(A)}_{z,\varepsilon}(t, v) &= \underline{a_v} t^2 + \phi_{z,\varepsilon}(v) - v^2 \underline{a_v} \\ &= \frac{z}{2} \phi_{z-2,\varepsilon}(v) t^2 + \underbrace{\phi_{z,\varepsilon}(v) - v^2 \frac{z}{2}, \phi_{z-2,\varepsilon}(v)}_{independent\ of\ t}. \end{aligned}$$

- If for every $v \in \mathbf{R}$ the parameter a_v is chosen independently of v as follows

$$a_v := \bar{a}_{z,\varepsilon}, \qquad \bar{a}_{z,\varepsilon} := \max_{v \in \mathbf{R}} \underline{a_v} = \max_{v \in \mathbf{R}} \left(\frac{\phi'_{z,\varepsilon}(v)}{2v} \right) = \frac{z}{2} \varepsilon^{z-2},$$

then the quadratic majorant takes the fixed form:

$$\begin{aligned} m^{(F)}_{z,\varepsilon}(t,v) &= \bar{a}_{z,\varepsilon} \left(t - v \left(1 - \underline{a_v} / \bar{a}_{z,\varepsilon} \right) \right)^2 \\ &+ \phi_{z,\varepsilon}(v) - v^2 \underline{a_v}^2 / \bar{a}_{z,\varepsilon} \\ &= \frac{z}{2} \left(\varepsilon^{z-2} t^2 - 2v \left(\varepsilon^{z-2} - \phi_{z-2,\varepsilon}(v) \right) t \right) \\ &+ \underbrace{\phi_{z,\varepsilon}(v) - v^2 \frac{z}{2} \left(2\phi_{z-2,\varepsilon}(v) - \varepsilon^{z-2} \right)}_{\textit{independent of } t} . \end{aligned}$$

Adaptive quadratic majorization

Replace the functions $\phi_{p,\varepsilon}$ and $\phi_{q,\varepsilon}$ in $\mathcal{J}_\varepsilon$ by associated adaptive quadratic majorants at $v_i^{(k)}$ and $u_j^{(k)}$. Gives adaptive quadratic majorant for $\mathcal{J}_\varepsilon(x)$ at $x^{(k)}$:

$$\begin{aligned}\mathcal{Q}^{(A)}(x, x^{(k)}) &= \frac{1}{2}\sum_{i=1}^{r} \phi_{p-2,\varepsilon}\big(v_i^{(k)}\big)(Ax-b)_i^2 \\ &+ \frac{\mu}{2}\sum_{j=1}^{s} \phi_{q-2,\varepsilon}\big(u_j^{(k)}\big)(Lx)_j^2 + c.\end{aligned}$$

The constant c is made up of terms that are independent of x.

Define vectors $w_{\text{fid}}^{(k)} \in \mathbf{R}^r$ and $w_{\text{reg}}^{(k)} \in \mathbf{R}^s$ of majorization weights for the fidelity and regularization terms:

$$\begin{aligned} w_{\text{fid}}^{(k)} &= \phi_{p-2,\varepsilon}\left(v^{(k)}\right) = \left(\left(v^{(k)}\right)^2 + \varepsilon^2\right)^{p/2-1}, \\ w_{\text{reg}}^{(k)} &= \phi_{q-2,\varepsilon}\left(u^{(k)}\right) = \left(\left(u^{(k)}\right)^2 + \varepsilon^2\right)^{q/2-1} \end{aligned}$$

and introduce the diagonal matrices

$$\begin{aligned} W_{\text{fid}}^{(k)} &= \operatorname{diag}\left(w_{\text{fid}}^{(k)}\right) \in \mathbf{R}^{r \times r}, \\ W_{\text{reg}}^{(k)} &= \operatorname{diag}\left(w_{\text{reg}}^{(k)}\right) \in \mathbf{R}^{s \times s}. \end{aligned}$$

Then

$$\begin{aligned}\mathcal{Q}^{(A)}(x, x^{(k)}) &= \frac{1}{2}\left\|\left(W_{\text{fid}}^{(k)}\right)^{1/2}(Ax - b)\right\|_2^2 \\ &+ \frac{\mu}{2}\left\|\left(W_{\text{reg}}^{(k)}\right)^{1/2} Lx\right\|_2^2 + c.\end{aligned}$$

Fixed quadratic majorization

Replace the functions $\phi_{p,\varepsilon}$ and $\phi_{q,\varepsilon}$ in $\mathcal{J}_\varepsilon$ by associated fixed quadratic majorants at $v_i^{(k)}$ and $u_j^{(k)}$. Gives fixed quadratic majorant for $\mathcal{J}_\varepsilon(x)$ at $x^{(k)}$:

$$\begin{aligned}\mathcal{Q}^{(F)}(x, x^{(k)}) &= \frac{\varepsilon^{p-2}}{2}\sum_{i=1}^{r}\left[(Ax-b)_i^2 - 2v_i^{(k)}\left(1-\frac{\phi_{p-2,\varepsilon}\big(v_i^{(k)}\big)}{\varepsilon^{p-2}}\right)(Ax-b)_i\right] \\ &+ \frac{\mu\varepsilon^{q-2}}{2}\sum_{j=1}^{s}\left[(Lx)_j^2 - 2u_j^{(k)}\left(1-\frac{\phi_{q-2,\varepsilon}\big(u_j^{(k)}\big)}{\varepsilon^{q-2}}\right)(Lx)_j\right] + c.\end{aligned}$$

Terms independent of x make up the constant c.

Define vectors $w_{\text{fid}}^{(k)} \in \mathbf{R}^r$ and $w_{\text{reg}}^{(k)} \in \mathbf{R}^s$ of majorization weights for the fidelity and regularization terms:

Component-wise

$$\begin{aligned} w_{\text{fid}}^{(k)} &= v^{(k)} \left(1 - \frac{\phi_{p-2,\varepsilon}\left(v^{(k)}\right)}{\varepsilon^{p-2}} \right), \\ w_{\text{reg}}^{(k)} &= u^{(k)} \left(1 - \frac{\phi_{q-2,\varepsilon}\left(u^{(k)}\right)}{\varepsilon^{q-2}} \right). \end{aligned}$$

The fixed quadratic majorant can be expressed in the compact form:

$$\begin{aligned}\mathcal{Q}^{(F)}(x, x^{(k)}) &= \frac{\varepsilon^{p-2}}{2}\left(\|Ax - b\|_2^2 - 2\left\langle w_{\text{fid}}^{(k)}, Ax\right\rangle\right) \\ &+ \frac{\mu\varepsilon^{q-2}}{2}\left(\|Lx\|_2^2 - 2\left\langle w_{\text{reg}}^{(k)}, Lx\right\rangle\right) + c.\end{aligned}$$

The minimization steps in the kth iteration of the adaptive MM approach can be written as

$$x^{(k+1)} = \arg \min_{x \in \mathbf{R}^n} \left[\left\| \left(W_{\text{fid}}^{(k)} \right)^{1/2} (Ax - b) \right\|_2^2 + \mu \left\| \left(W_{\text{reg}}^{(k)} \right)^{1/2} Lx \right\|_2^2 \right]$$

and of the fixed MM approach as

$$x^{(k+1)} = \arg \min_{x \in \mathbf{R}^n} \left[\|Ax - b\|_2^2 - 2 \left\langle w_{\text{fid}}^{(k)}, Ax \right\rangle + \eta \left(\|Lx\|_2^2 - 2 \left\langle w_{\text{reg}}^{(k)}, Lx \right\rangle \right) \right].$$

Terms independent of x are omitted and $\eta := \mu \frac{\varepsilon^{q-2}}{\varepsilon^{p-2}}$.

Define the $n \times n$ matrices

$$\begin{aligned} T^{(A)}\left(W_{\text{fid}}, W_{\text{reg}}\right) \quad &:= \quad A^T W_{\text{fid}} A \,+\, \mu\, L^T W_{\text{reg}} L, \\ T^{(F)} \quad &:= \quad A^T A \,+\, \eta\, L^T L\ . \end{aligned}$$

The normal equations associated with the adaptive and fixed quadratic minimization problems can be written

$$\begin{aligned} T^{(A)}\left(W_{\text{fid}}^{(k)}, W_{\text{reg}}^{(k)}\right) x \quad &= \quad A^T W_{\text{fid}}^{(k)} b\ , \\ T^{(F)}\, x \quad &= \quad A^T \left(b + w_{\text{fid}}^{(k)}\right) + \eta\, L^T w_{\text{reg}}^{(k)}\ . \end{aligned}$$

The normal equations for the adaptive approach have a unique solution if

$$\mathrm{Ker}\left(A^T W_{\mathrm{fid}}^{(k)} A\right) \cap \mathrm{Ker}\left(L^T W_{\mathrm{reg}}^{(k)} L\right) = \{0\} \quad \forall k,$$

and the normal equations for the fixed approach have a unique solution if

$$\mathrm{Ker}\left(A^T A\right) \cap \mathrm{Ker}\left(L^T L\right) = \{0\},$$

The generalized Krylov subspace (GKS) method

Let the columns of $V_k \in \mathbf{R}^{n\times k}$ form an orthonormal basis for the (generalized Krylov) solution subspace.

Adaptive minimization problem:

$$\min_{y\in\mathbf{R}^{k+l}} \left\| \begin{bmatrix} \left(W_{\text{fid}}\right)^{1/2} A V_k \\ \mu^{1/2}\left(W_{\text{reg}}\right)^{1/2} L V_k \end{bmatrix} y - \begin{bmatrix} \left(W_{\text{fid}}\right)^{1/2} b \\ 0 \end{bmatrix} \right\|_2^2 .$$

Fixed minimization problem:

$$\min_{y\in\mathbf{R}^{k+l}} \left\| \begin{bmatrix} A V_k \\ \eta^{1/2} L V_k \end{bmatrix} y - \begin{bmatrix} b + w_{\text{fid}}^{(k)} \\ \eta^{1/2} w_{\text{reg}}^{(k)} \end{bmatrix} \right\|_2^2 .$$

Solution $y^{(k+1)}$. Then $x^{(k+1)} := V_k y^{(k+1)}$.

The GKS method for the fixed minimization problem

Let $V_k \in \mathbf{R}^{n\times d}$, $d = k + l \ll n$. Define the QR factorizations

$$
\begin{aligned}
AV_k &= Q_A R_A \quad \text{with} \quad Q_A \in \mathbf{R}^{r\times d}, \quad R_A \in \mathbf{R}^{d\times d}\,, \\
LV_k &= Q_L R_L \quad \text{with} \quad Q_L \in \mathbf{R}^{s\times d}\,, \quad R_L \in \mathbf{R}^{d\times d}\,.
\end{aligned}
$$

Substituting factorizations into the minimization problem gives small problem:

$$
\min_{y\in\mathbf{R}^{k+l}} \left\| \begin{bmatrix} R_A \\ \eta^{1/2} R_L \end{bmatrix} y - \begin{bmatrix} Q_A^T (b + w_{\text{fid}}^{(k)}) \\ \eta^{1/2} Q_L^T w_{\text{reg}}^{(k)} \end{bmatrix} \right\|_2^2 .
$$

Residual vector for the normal equations:

$$\begin{aligned} r^{(k+1)} &= T^{(F)}\,x^{(k+1)} - A^T\left(b + w_{\text{fid}}^{(k)}\right) - \eta L^T w_{\text{reg}}^{(k)} \\ &= A^T\big(AV_k y^{(k+1)} - b - w_{\text{fid}}^{(k)}\big) + \eta L^T\big(LV_k y^{(k+1)} - w_{\text{reg}}^{(k)}\big). \end{aligned}$$

The subspace $\mathcal{V}_k$ is expanded to $\mathcal{V}_{k+1}$ by adding the new basis vector

$$v_{\text{new}} := \frac{r^{(k+1)}}{\|r^{(k+1)}\|_2}, \qquad V_{k+1} := [V_k, v_{\text{new}}]\,.$$

To enforce orthogonality in the presence of round-off errors, v_{new} is reorthogonalized against V_k.

Update the QR factorizations

$$
\begin{aligned}
A[V_k, v_{\text{new}}] &= [Q_A, \tilde{q}_{A,k+1}] \begin{bmatrix} R_A & r_{K,k+1} \\ 0 & \tau_{K,k+1} \end{bmatrix}, \\
L[V_k, v_{\text{new}}] &= [Q_L, \tilde{q}_{L,k+1}] \begin{bmatrix} R_L & r_{L,k+1} \\ 0 & \tau_{L,k+1} \end{bmatrix}.
\end{aligned}
$$

The GKS method for the adaptive minimization problem

Let $V_k \in \mathbf{R}^{n\times d}$, $d = k + l \ll n$. Compute the QR factorizations

$$W_{\rm fid}^{1/2} A V_k \quad = \quad Q_A R_A \quad \text{with} \quad Q_A \in \mathbf{R}^{r\times d}, \quad R_A \in \mathbf{R}^{d\times d}\,,$$
$$W_{\rm reg}^{1/2} L V_k \quad = \quad Q_L R_L \quad \text{with} \quad Q_L \in \mathbf{R}^{s\times d}\,, \quad R_L \in \mathbf{R}^{d\times d}\,.$$

Substituting into minimization problem gives small problem

$$\min_{y\in\mathbf{R}^{k+l}} \left\| \begin{bmatrix} R_A \\ \mu^{1/2} R_L \end{bmatrix} y - \begin{bmatrix} Q_A^T W_{\rm fid}^{1/2} b \\ 0 \end{bmatrix} \right\|_2^2 .$$

Residual vector for the normal equations:

$$\begin{aligned} r^{(k+1)} &= T\big(W_{\text{fid}}, W_{\text{reg}}\big)\, x^{(k+1)} - A^T W_{\text{fid}} b \\ &= A^T W_{\text{fid}}\big(A V_k y^{(k+1)} - b\big) + \mu L^T W_{\text{reg}}\big(L V_k y^{(k+1)}\big). \end{aligned}$$

is normalized, reorthogonalized, and appended to the matrix V_k.

Convergence analysis for the MM-GKS methods

The MM-GKS approach can be written in the form

$$x^{(k+1)} := \begin{cases} \arg\min\limits_{x\in\mathcal{V}_k} Q(x, x^{(k)}) & \text{for } \; k = 0, 1, \ldots, n-l-1, \\ \arg\min\limits_{x\in\mathbf{R}^n} Q(x, x^{(k)}) & \text{for } \; k = n-l, n-l+1, \ldots \; , \end{cases}$$

where

- $l \geq 1$ is the dimension of the user-specified initial subspace $\mathcal{V}_0$,
- $\mathcal{V}_k$ is the generalized Krylov subspace used at iteration k,
- $Q(x, x^{(k)})$ is either $Q^{(A)}(x, x^{(k)})$ or $Q^{(F)}(x, x^{(k)})$.

Theorem: Let each linearized minimization problem have a unique solution. Then, for any initial guess $x^{(0)} \in \mathbf{R}^n$, the sequence $\{\mathcal{J}_\varepsilon(x^{(k)})\}_{k \geq 0}$ is monotonically non-increasing and convergent.

Theorem: Let each linearized minimization problem have a unique solution. Then, for any initial guess $x^{(0)} \in \mathbf{R}^n$, the sequence $\{x^{(k)}\}_{k\geq 1}$ converges to a stationary point of $\mathcal{J}_\varepsilon(x)$. Thus,

a. $$\lim_{k\to\infty} \left\| x^{(k+1)} - x^{(k)} \right\|_2 = 0,$$

b. $$\lim_{k\to\infty} \nabla_x \mathcal{J}_\varepsilon\left(x^{(k)}\right) = 0.$$

Corollary: If in addition $p > 1$ and $q > 1$, then for any initial guess $x^{(0)} \in \mathbf{R}^n$, the sequence $\left\{x^{(k)}\right\}_{k\geq 0}$ converges towards the unique global minimizer of the smoothed ℓ_p-ℓ_q functional.

Determining the regularization parameter for $p = 2$ by the discrepancy principle

Assume that a bound $\|b - b_{\text{exact}}\|_2 \leq \delta$ is known.

- Use a monotonically decreasing sequence of regularization parameter values $\mu = \mu_k$. Let $x^{(k)}$ be the solution of the minimization problem with $\mu = \mu_k$ and assume that the matrix A is nonsingular. Terminate the above iterations with the discrepancy principle, i.e., as soon as $\|Ax^{(k)} - b\|_2 \leq \delta$, Then

$$\limsup_{\delta \searrow 0} \|x^{(k)} - x_{\text{exact}}\|_2 = 0.$$

- Choose $\mu = \mu_k$ in each iteration so that $\|Ax^{(k)} - b\|_2 = \delta$. Then there is a subsequence $x^{(k_j)}$, $j = 1, 2, \ldots$, of computed solutions such that

$$\lim_{\delta \searrow 0} \sup \|x^{(k_j)} - x_{\text{exact}}\|_2 = 0.$$

Determining the regularization parameter by (standard) cross validation

Consider for simplicity Tikhonov regularization in standard form

$$\min_{x \in \mathbf{R}^n} \{ \|Ax - b\|_2^2 + \mu \|x\|_2^2 \}.$$

The CV method partitions b into two subsets (several times): the training set and the testing set.

- The training set is used for solving the problem (with the rows of the testing set removed) for different regularization parameters.

- The testing set is used to validate the computed solution and select a suitable regularization parameter.

Assume first that the testing set consists of the first d rows of A and b. Let

$$\widetilde{b} = [b_{d+1}, b_{d+2}, \ldots, b_m]^T,$$

$$\widetilde{A} = \begin{bmatrix} A_{d+1,1} & A_{d+1,2} & \cdots & A_{d+1,n} \\ A_{d+2,1} & A_{d+2,2} & \cdots & A_{d+2,n} \\ \vdots & \vdots & \ldots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix}.$$

Let $\mu_1 > \mu_2 > \ldots > \mu_l > 0$ be regularization parameters, say,

$$\mu_{j+1} = \mu_j/10, \qquad j = 1, 2, \ldots, l-1.$$

Solve Tikhonov minimization problem with A, b, and μ replaced by $\widetilde{A}$, $\widetilde{b}$, and μ_j, repectively. Gives solutions x_{μ_j}, $j = 1, 2, \ldots, l$.

Validate x_{μ_j} with the testing set, i.e., compute

$$\rho_j = \sqrt{\sum_{i=1}^{d} \left(\left(A x_{\mu_j} \right)_i - b_i \right)^2}, \quad j = 1, 2, \ldots, l.$$

Let $1 \leq j^* \leq l$ be such that

$$\rho_{j^*} \leq \rho_j, \quad j = 1, 2, \ldots, l.$$

Let $\mu = \mu_{j^*}$.

Repeat for new training set obtained by removing d other rows of A and d. In the computed examples $d = l = 10$ and 10 training sets.

Choose the regularization parameter to be the average of the μ-values computed.

Algorithm

for $k = 1, 2, \ldots, K$

$\widetilde{A}$ and $\widetilde{b}$ versions of A and b, in which the kth set of d consecutive rows have been removed

for $j = 1, 2, \ldots, l$

Let $x_{\mu_j}^{(k)}$ be Tikhonov solution with A, b, and μ replaced by $\widetilde{A}$, $\widetilde{b}$, and μ_j

$$r_j^{(k)} = \sqrt{\sum_{i=d(k-1)+1}^{kd} \left(\left(Ax_{\mu_j}^{(k)}\right)_i - b_i\right)^2}$$

$j^* = \arg\min_{1 \le j \le l}\{r_j^{(k)}\}$

$\mu^{(k)} = \mu_{j^*}$

end

$\mu = \frac{1}{K}\sum_{k=1}^{K} \mu^{(k)}$

Determining the regularization parameter by modified cross validation

Compare predictions of computed solutions:

- Let I_1 and I_2 be distinct sets of d distinct random integers in $[1, n]$.
- For $i = 1, 2$, let $\widetilde{A}_i$ and $\widetilde{b}_i$ be obtained by removing rows with indices in I_i from A and b.
- Let $\mu_1 > \mu_2 > \ldots > \mu_l > 0$ be regularization parameters.

- For $i = 1, 2$, let $x_{\mu_j}^{(i)}$ solve the Tikhonov regularization problem A, b, and μ replaced by $\widetilde{A}_i$, $\widetilde{b}$, and μ_j.

- Compute

$$\Delta x_j = \|x_{\mu_j}^{(1)} - x_{\mu_j}^{(2)}\|_2, \qquad j = 1, 2, \ldots, l,$$

- Let $\mu^{(k)}$ minimize Δx_j over $j = 1, 2, \ldots, l$.

Repeat for several sets I_1 and I_2. Let μ be the average of the $\mu^{(k)}$ computed.

The IRN method

Method proposed by Rodriguez and Wohlberg. Apply the conjugate gradient method to solve the sequence normal equations determined by adaptive approach.

Computed examples

We report the Signal-to-Noise ratio (SNR):

$$\mathrm{SNR}(x^*, \bar{x}) := 10 \log_{10} \frac{\|\bar{x} - E(\bar{x})\|_2^2}{\|x^* - \bar{x}\|_2^2} \ \text{(dB)}\ ,$$

where $E(\bar{x})$ denotes the mean gray level of the uncontaminated image $\bar{x}$.

We use the initial subspace

$$\mathcal{V}_0 = \mathrm{span}\{A^T b\}. \qquad (l = 1)$$

Example: Cameraman image: 512×512 pixels.

Original image

Contaminated image, SNR=-0.80

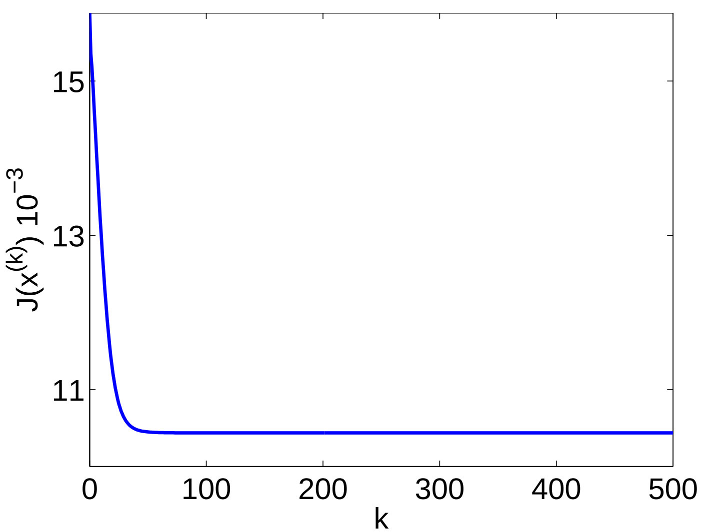
20% salt-and-pepper noise, Gaussian blur.

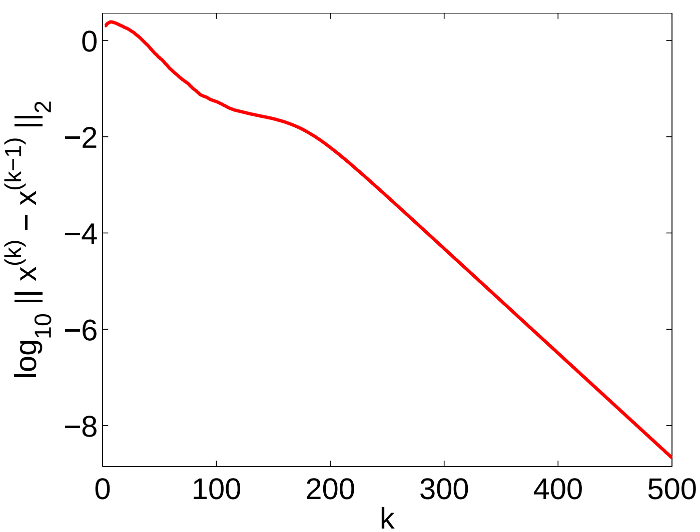
Restored image by ℓ_1-ℓ_1 TV minimization, SNR=13.22.

Restored image by $\ell_{0.7}$-$\ell_{1.0}$ TV minimization, SNR=15.33.

blur		noise		efficiency: time (iterations, MVPs)			accuracy: SNR		
band	σ	%	μ	IRN	AMM	FMM	IRN	AMM	FMM
ℓ_1-ℓ_1									
9	2.5	10	0.004	303.12 (39,6182)	163.08 (177,708)	42.83 (202,808)	13.00	12.99	12.98
		20	0.005	291.40 (42,5892)	155.69 (174,696)	44.63 (203,812)	12.01	12.01	12.05
		30	0.020	180.29 (55,3586)	65.89 (123,492)	29.23 (162,648)	11.64	11.65	11.69
$\ell_{0.7}$-ℓ_1									
9	2.5	10	0.004	497.87 (34,10292)	427.96 (256,1024)	70.49 (274,1096)	15.20	15.19	15.15
		20	0.006	430.04 (37,8838)	300.07 (224,896)	69.08 (265,1060)	14.29	14.28	14.26
		30	0.010	365.49 (41,7450)	224.34 (201,804)	67.20 (266,1064)	13.47	13.47	13.43

Convergence of functional and difference between consecutive iterates.





Application of wavelets

Compute a sparse solution by using a two-level framelet analysis operator as regularization operator L. Framelets are extensions of wavelets.

Let $\mathbb{A} \in \mathbf{R}^{r \times n}$ with $n \leq r$. The set of the rows of $\mathbb{A}$ is a tight frame for $\mathbb{R}^n$ if

$$\|x\|_2^2 = \sum_{j=1}^{r} y_j^T x \quad \forall\, x \in \mathbf{R}^n,$$

where $y_j \in \mathbf{R}^n$ is the jth row of $\mathbb{A}$ (written as a column vector), i.e., $\mathbb{A} = [y_1, \ldots, y_r]^T$. The matrix $\mathbb{A}$ is referred to as an analysis operator and $\mathbb{A}^T$ as a synthesis operator.

Tight frames determined by B-splines: Made up of a low-pass filter W_0 and two high-pass filters W_1 and W_2 defined by the masks

$$w^{(0)} = \frac{1}{2}\,(1,\,2,\,1)\,,\; w^{(1)} = \frac{\sqrt{2}}{4}\,(1,\,0,\,-1)\,,\; w^{(2)} = \frac{1}{4}\,(-1,\,2,\,-1)\,.$$

These masks and reflective boundary conditions yield the matrices

$$W_0 = \frac{1}{4}\begin{pmatrix} 3 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 2 & 1 \\ 0 & \dots & 0 & 1 & 3 \end{pmatrix},\; W_1 = \frac{\sqrt{2}}{4}\begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix},$$

$$W_2 = \frac{1}{4}\begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & \dots & 0 & 1 & 1 \end{pmatrix}.$$

This gives the analysis operator for problems in 1D:

$$\mathbb{A} = \begin{bmatrix} W_0 \\ W_1 \\ W_2 \end{bmatrix}$$

with $\mathbb{A}^T\mathbb{A} = I$.

In 2D: Let $W_{ij} = W_i \otimes W_j$ and define

$$\mathbb{A} = \begin{bmatrix} W_{00} \\ W_{01} \\ \vdots \\ W_{22} \end{bmatrix}.$$

Original image, 246 × 246 pixels.

PSF, 9×9 pixels.

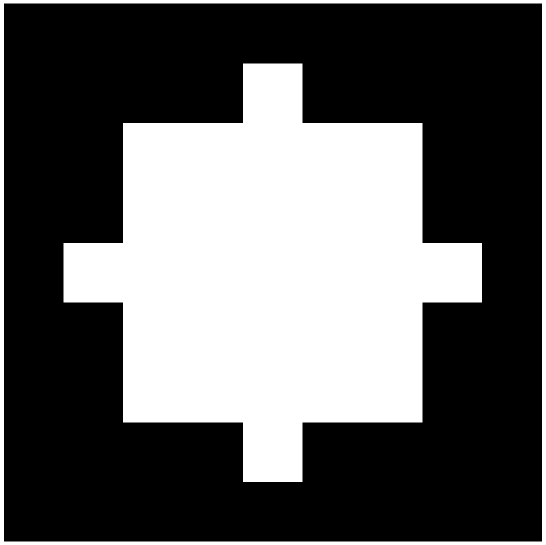


Image contaminated by blur and 1% white Gaussian noise.

Restored image by FMM-GKS, best regularization parameter.

Restored image by FMM-GKS, regularization parameter determined by monotonically decreasing sequence.

Restored image by FMM-GKS, regularization parameter determined by discrepancy principle.

Restoration of clock image.

Method	relative error	# iterations
MM-GKS-R	0.032887	29
MM-GKS-MD	0.036663	5
MM-GKS-DP	**0.032828**	28

New PSF, 29×29 pixels.

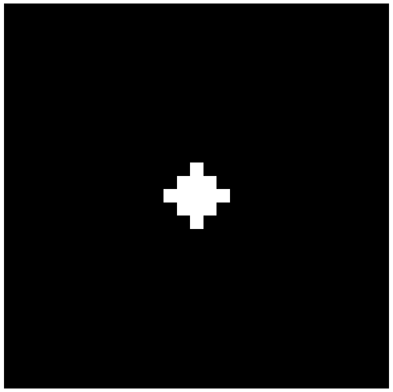


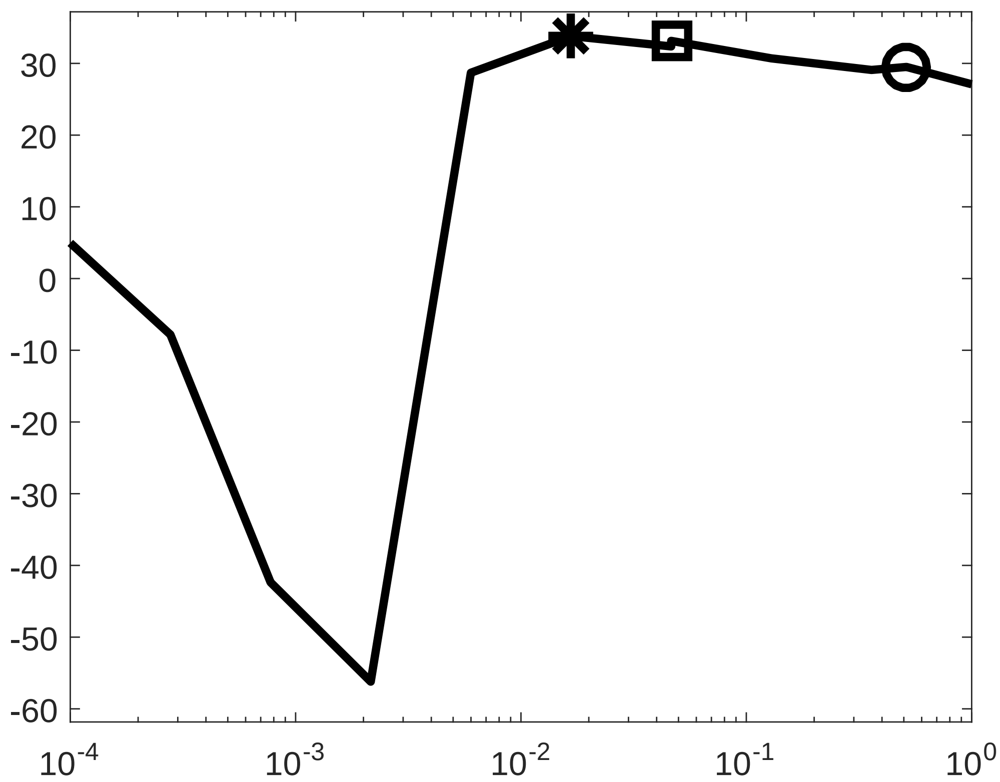
Image contaminated by blur and 10% salt-and pepper noise.

Restoration, optimal regularization parameter:
PSNR=33.81.

Restoration, regularization parameter by cross validation: PSNR=29.51.

Restoration, regularization parameter by modified cross validation: PSNR=33.14.

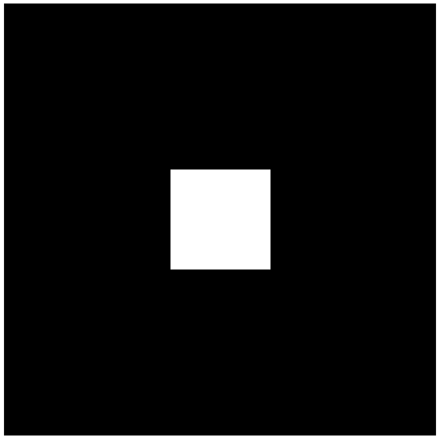
PSNR vs. the regularization parameter: star=optimal, square=MCV, circle=CV.



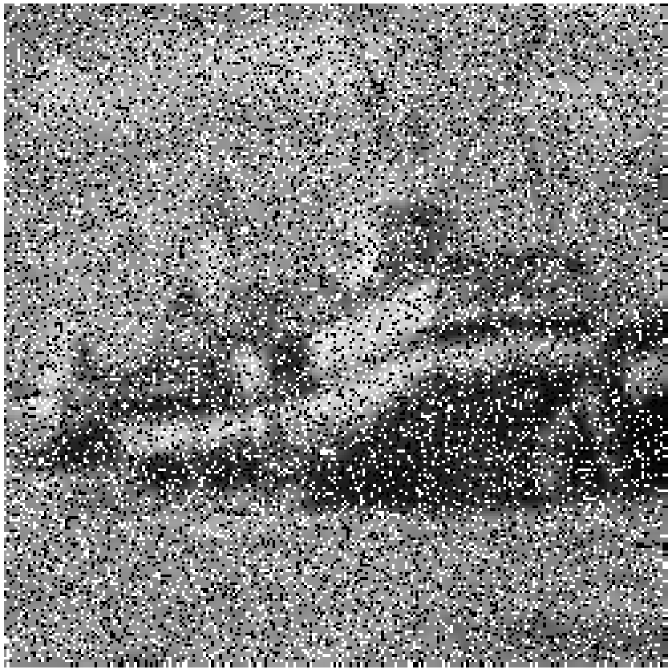
250×250 pixels

Original image

Out-of-focus PSF, 26×26 pixels



Blurred and noisy image, 30% salt-and-pepper noise



Restored image, optimal μ, $p = 0.8$, $q = 0.5$,
PSNR$= 26.15$

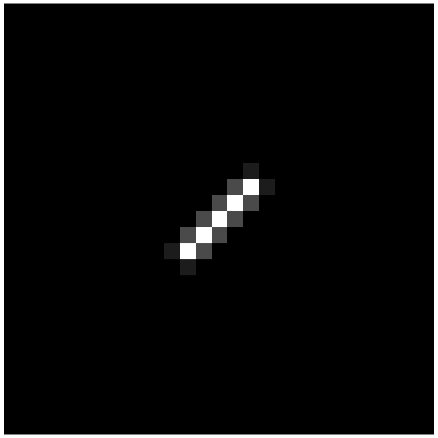
Restored image, CV μ, $p = 0.8$, $q = 0.5$, PSNR$= 25.27$

Restored image, MCV μ, $p = 0.8$, $q = 0.5$, PSNR$= 26.15$

247×247 pixels

Original image

Motion PSF, 27×27 pixels



Blurred and noisy image, 10% salt-and-pepper noise and white Gaussian noise

Restored image, optimal μ, $p = 0.8$, $q = 0.1$, PSNR$= 26.58$

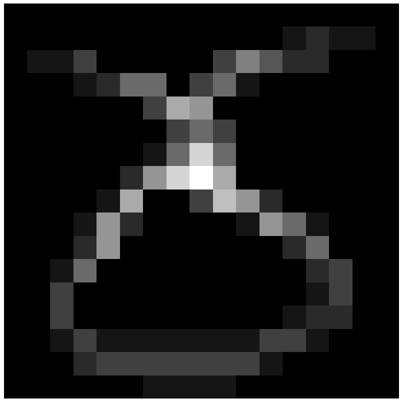
Restored image, CV μ, $p = 0.8$, $q = 0.1$, PSNR$= 25.49$

Restored image, MCV μ, $p = 0.8$, $q = 0.1$, PSNR$= 25.74$

234×182 pixels

Original image

Motion PSF, 17×17 pixels



Blurred and noisy image, 20% salt-and-pepper noise and 1% Gaussian noise

Restored image, CV, $p = 0.8$, $q = 0.1$, relative error$= 0.0740$

Restored image, MCV, $p = 0.8$, $q = 0.1$, relative error= 0.0689

Grazie 10^3