

Image restoration: numerical optimisation

— Short and partial presentation —

Jean-François Giovannelli

Groupe Signal – Image

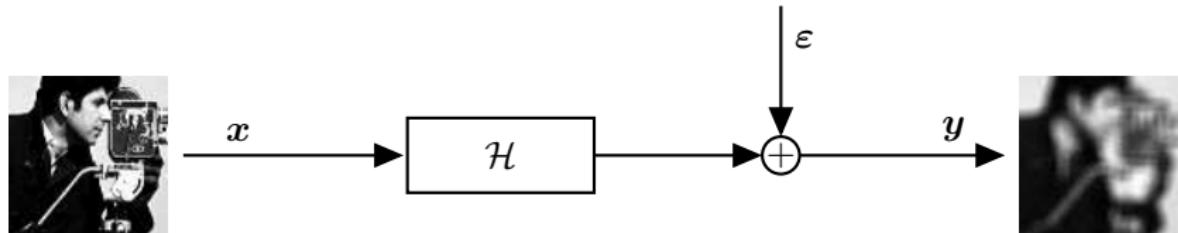
Laboratoire de l'Intégration du Matériau au Système
Univ. Bordeaux – CNRS – BINP

Context and topic

- Image restoration, deconvolution
 - Missing information: ill-posedness
 - Compensation: regularization
- Previous / next lectures, exercises and practical works
 - Three types of regularised inversion
 - Quadratic penalties and smoothness
 - Convex non-quadratic penalties and edge preservation
 - Constraints: positivity and support
 - Bayesian strategy: an incursion for hyperparameter tuning
- A basic component: Quadratic criterion and Gaussian model
 - Circulant approximation and computations based on FFT
 - Other approaches:
 - Numerical linear system solvers
 - Numerical quadratic optimizers

Direct / Inverse, Convolution / Deconvolution, . . .

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\varepsilon} = \mathbf{h} \star \mathbf{x} + \boldsymbol{\varepsilon}$$



$$\hat{\mathbf{x}} = \hat{\mathcal{X}}(\mathbf{y})$$

Reconstruction, Restoration, Deconvolution, Denoising

- General problem: ill-posed inverse problems, *i.e.*, *lack of information*
- Methodology: regularisation, *i.e.*, *information compensation*

Various quadratic criteria

- Quadratic criterion and linear solution

$$\mathcal{J}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu \sum (\mathbf{x}_p - \mathbf{x}_q)^2 = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu \|\mathbf{D}\mathbf{x}\|^2$$

- Huber penalty, extended criterion and half-quadratic approaches

$$\mathcal{J}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu \sum \varphi(\mathbf{x}_p - \mathbf{x}_q)$$

$$\tilde{\mathcal{J}}(\mathbf{x}, \mathbf{a}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu \sum \frac{1}{2} [(\mathbf{x}_p - \mathbf{x}_q) - \mathbf{a}_{pq}]^2 + \zeta(\mathbf{a}_{pq})$$

- Constraints: augmented Lagrangian and ADMM

$$\begin{cases} \mathcal{J}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu \|\mathbf{D}\mathbf{x}\|^2 \\ \text{s.t. } \begin{cases} \mathbf{x}_p = 0 & \text{for } p \in \bar{\mathcal{S}} \\ \mathbf{x}_p \geq 0 & \text{for } p \in \mathcal{M} \end{cases} \end{cases}$$

$$\mathcal{L}(\mathbf{x}, \mathbf{s}, \boldsymbol{\ell}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu \|\mathbf{D}\mathbf{x}\|^2 + \rho \|\mathbf{x} - \mathbf{s}\|^2 + \boldsymbol{\ell}^t(\mathbf{x} - \mathbf{s})$$

Various solutions (no circulant approximation)...



True



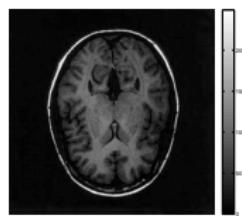
Observation



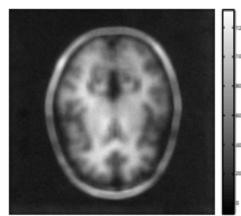
Quadratic penalty



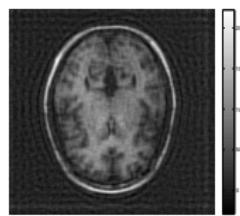
Huber penalty



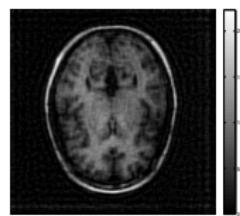
True



Observation



Quadratic penalty



Constrained

A reminder of the calculi

- The criterion...

$$\mathcal{J}(\mathbf{x}) = \|\mathbf{y} - \mathbf{Hx}\|^2 + \mu \|\mathbf{Dx}\|^2$$

- ... its gradient ...

$$\begin{aligned}\frac{\partial \mathcal{J}}{\partial \mathbf{x}} &= -2\mathbf{H}^t(\mathbf{y} - \mathbf{Hx}) + 2\mu \mathbf{D}^t \mathbf{Dx} \\ &= 2(\mathbf{H}^t \mathbf{H} + \mu \mathbf{D}^t \mathbf{D})\mathbf{x} - 2\mathbf{H}^t \mathbf{y}\end{aligned}$$

- ... and its Hessian ...

$$\frac{\partial^2 \mathcal{J}}{\partial \mathbf{x}^2} = 2(\mathbf{H}^t \mathbf{H} + \mu \mathbf{D}^t \mathbf{D})$$

- The multivariate linear system of equations...

$$(\mathbf{H}^t \mathbf{H} + \mu \mathbf{D}^t \mathbf{D}) \bar{\mathbf{x}} = \mathbf{H}^t \mathbf{y}$$

- ... and the minimiser...

$$\bar{\mathbf{x}} = (\mathbf{H}^t \mathbf{H} + \mu \mathbf{D}^t \mathbf{D})^{-1} \mathbf{H}^t \mathbf{y}$$

Reformulation, notations . . .

- Rewrite the criterion . . .

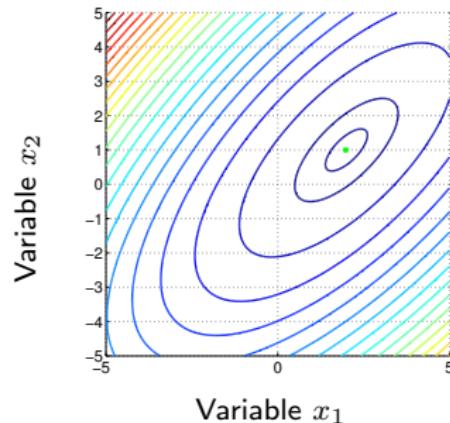
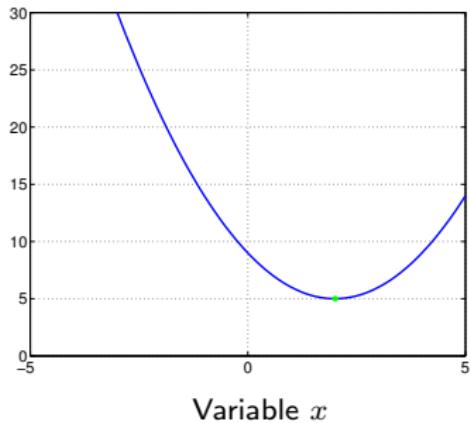
$$\begin{aligned}\mathcal{J}(\mathbf{x}) &= \|\mathbf{y} - \mathbf{Hx}\|^2 + \mu \|\mathbf{Dx}\|^2 \\ &= \frac{1}{2} \mathbf{x}^t \mathbf{Qx} + \mathbf{q}^t \mathbf{x} + q_0\end{aligned}$$

- Gradient: $\mathbf{g}(\mathbf{x}) = \mathbf{Qx} + \mathbf{q} = 2(\mathbf{H}^t \mathbf{H} + \mu \mathbf{D}^t \mathbf{D})\mathbf{x} - 2\mathbf{H}^t \mathbf{y}$
- Hessian: $\mathbf{Q} = 2(\mathbf{H}^t \mathbf{H} + \mu \mathbf{D}^t \mathbf{D})$
- $\mathbf{q} = \mathbf{g}(\mathbf{0}) = -2\mathbf{H}^t \mathbf{y}$ the gradient at $\mathbf{x} = 0$
- $q_0 = \mathcal{J}(\mathbf{0}) = \|\mathbf{y}\|^2$
- . . . the system of equations . . .
$$\begin{aligned}(\mathbf{H}^t \mathbf{H} + \mu \mathbf{D}^t \mathbf{D}) \bar{\mathbf{x}} &= \mathbf{H}^t \mathbf{y} \\ \mathbf{Q} \bar{\mathbf{x}} &= -\mathbf{q}\end{aligned}$$
- . . . and the minimiser . . .
$$\begin{aligned}\bar{\mathbf{x}} &= (\mathbf{H}^t \mathbf{H} + \mu \mathbf{D}^t \mathbf{D})^{-1} \mathbf{H}^t \mathbf{y} \\ &= -\mathbf{Q}^{-1} \mathbf{q}\end{aligned}$$

Quadratic criterion: illustration

① $\alpha^2 (x - \bar{x})^2 + \gamma$

② $\alpha_1^2 (x_1 - \bar{x}_1)^2 + \alpha_2^2 (x_2 - \bar{x}_2)^2 - 2\rho\alpha_1\alpha_2(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) + \gamma$

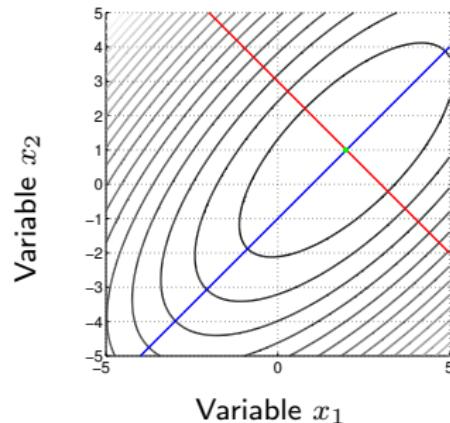
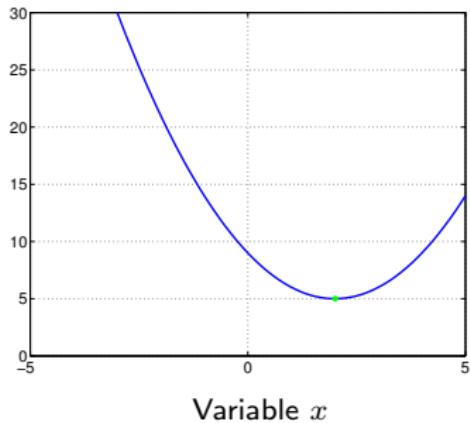


- Comment: convex, concave, saddle point, proper direction

Quadratic criterion: illustration

① $\alpha^2 (x - \bar{x})^2 + \gamma$

② $\alpha_1^2 (x_1 - \bar{x}_1)^2 + \alpha_2^2 (x_2 - \bar{x}_2)^2 - 2\rho\alpha_1\alpha_2(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) + \gamma$



- Comment: convex, concave, saddle point, proper direction

Computational and algorithmic aspects

Various options and many relationships...

- Direct calculus, compact (closed) form, matrix inversion
- Circulant approximation and diagonalisation by FFT
- Special algorithms for 1D case
 - Recursive least squares
 - Kalman smoother or filter (and fast versions, ...)
- Algorithms for **linear system solving**
 - **Splitting idea**
 - Gauss, Gauss-Jordan
 - Substitution, triangularisation, ...
 - Et Levinson, ...
- Algorithms for **criterion optimisation**
 - **Component wise...**
 - **Gradient descent...**
 - ... and various modifications

A family of **linear system solvers**

Matrix splitting algorithms

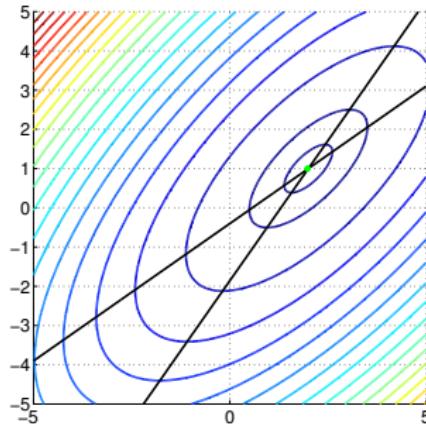
Linear solver: illustration

- Criterion

$$\mathcal{J}(\boldsymbol{x}) = \alpha_1^2 (x_1 - \bar{x}_1)^2 + \alpha_2^2 (x_2 - \bar{x}_2)^2 - 2\rho\alpha_1\alpha_2(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) + \gamma$$

- Nullification of the gradient

$$\begin{cases} \partial \mathcal{J} / \partial x_1 = 2\alpha_1^2 (x_1 - \bar{x}_1) - 2\rho\alpha_1\alpha_2(x_2 - \bar{x}_2) = 0 \\ \partial \mathcal{J} / \partial x_2 = 2\alpha_2^2 (x_2 - \bar{x}_2) - 2\rho\alpha_1\alpha_2(x_1 - \bar{x}_1) = 0 \end{cases}$$



An example of linear solver: matrix splitting

A simple and fruitful idea

- Solve $\mathbf{Q} \mathbf{x} + \mathbf{q} = \mathbf{0}$ that is $\mathbf{Q} \mathbf{x} = -\mathbf{q}$
- Matrix splitting $\mathbf{Q} = \mathbf{A} - \mathbf{B}$
- Take advantage...

$$\begin{aligned}\mathbf{Q} \mathbf{x} &= -\mathbf{q} \\ (\mathbf{A} - \mathbf{B}) \mathbf{x} &= -\mathbf{q} \\ \mathbf{A} \mathbf{x} &= \mathbf{B} \mathbf{x} - \mathbf{q} \\ \mathbf{x} &= \mathbf{A}^{-1} [\mathbf{B} \mathbf{x} - \mathbf{q}]\end{aligned}$$

An example of linear solver: matrix splitting

A simple and fruitful idea

- Solve $\mathbf{Q} \mathbf{x} + \mathbf{q} = \mathbf{0}$ that is $\mathbf{Q} \mathbf{x} = -\mathbf{q}$
- Matrix splitting $\mathbf{Q} = \mathbf{A} - \mathbf{B}$
- Take advantage...

$$\begin{aligned}\mathbf{Q} \mathbf{x} &= -\mathbf{q} \\ (\mathbf{A} - \mathbf{B}) \mathbf{x} &= -\mathbf{q} \\ \mathbf{A} \mathbf{x} &= \mathbf{B} \mathbf{x} - \mathbf{q} \\ \mathbf{x} &= \mathbf{A}^{-1} [\mathbf{B} \mathbf{x} - \mathbf{q}]\end{aligned}$$

- Iterative solver

An example of linear solver: matrix splitting

A simple and fruitful idea

- Solve $\mathbf{Q} \mathbf{x} + \mathbf{q} = \mathbf{0}$ that is $\mathbf{Q} \mathbf{x} = -\mathbf{q}$
- Matrix splitting $\mathbf{Q} = \mathbf{A} - \mathbf{B}$
- Take advantage...

$$\begin{aligned}\mathbf{Q} \mathbf{x} &= -\mathbf{q} \\ (\mathbf{A} - \mathbf{B}) \mathbf{x} &= -\mathbf{q} \\ \mathbf{A} \mathbf{x} &= \mathbf{B} \mathbf{x} - \mathbf{q} \\ \mathbf{x} &= \mathbf{A}^{-1} [\mathbf{B} \mathbf{x} - \mathbf{q}]\end{aligned}$$

- Iterative solver

$$\mathbf{x}^{[k]} = \mathbf{A}^{-1} [\mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{q}]$$

- If it exists, a fixed point satisfies:

$$\begin{aligned}\mathbf{x}^{\infty} &= \mathbf{A}^{-1} [\mathbf{B} \mathbf{x}^{\infty} - \mathbf{q}] \\ \text{so, } \mathbf{Q} \mathbf{x}^{\infty} &= -\mathbf{q}\end{aligned}$$

Matrix splitting: sketch of proof (1)

- Notations

$$\begin{aligned}\boldsymbol{x}^{[k]} &= \boldsymbol{A}^{-1} [\boldsymbol{B} \boldsymbol{x}^{[k-1]} - \boldsymbol{q}] \\ &= \boldsymbol{A}^{-1} \boldsymbol{B} \boldsymbol{x}^{[k-1]} - \boldsymbol{A}^{-1} \boldsymbol{q} \\ &= \boldsymbol{M} \boldsymbol{x}^{[k-1]} - \boldsymbol{m}\end{aligned}$$

- Iterates

Matrix splitting: sketch of proof (1)

- Notations

$$\begin{aligned}\mathbf{x}^{[k]} &= \mathbf{A}^{-1} [\mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{q}] \\ &= \mathbf{A}^{-1} \mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{A}^{-1} \mathbf{q} \\ &= \mathbf{M} \mathbf{x}^{[k-1]} - \mathbf{m}\end{aligned}$$

- Iterates

$$\mathbf{x}^{[K]} = \mathbf{M} \mathbf{x}^{[K-1]} - \mathbf{m}$$

Matrix splitting: sketch of proof (1)

- Notations

$$\begin{aligned}\mathbf{x}^{[k]} &= \mathbf{A}^{-1} [\mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{q}] \\ &= \mathbf{A}^{-1} \mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{A}^{-1} \mathbf{q} \\ &= \mathbf{M} \mathbf{x}^{[k-1]} - \mathbf{m}\end{aligned}$$

- Iterates

$$\begin{aligned}\mathbf{x}^{[K]} &= \mathbf{M} \mathbf{x}^{[K-1]} - \mathbf{m} \\ &= \mathbf{M} (\mathbf{M} \mathbf{x}^{[K-2]} - \mathbf{m}) - \mathbf{m} \\ &= \mathbf{M}^2 \mathbf{x}^{[K-2]} - (\mathbf{M} + \mathbf{I}) \mathbf{m}\end{aligned}$$

Matrix splitting: sketch of proof (1)

- Notations

$$\begin{aligned}\mathbf{x}^{[k]} &= \mathbf{A}^{-1} [\mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{q}] \\ &= \mathbf{A}^{-1} \mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{A}^{-1} \mathbf{q} \\ &= \mathbf{M} \mathbf{x}^{[k-1]} - \mathbf{m}\end{aligned}$$

- Iterates

$$\begin{aligned}\mathbf{x}^{[K]} &= \mathbf{M} \mathbf{x}^{[K-1]} - \mathbf{m} \\ &= \mathbf{M} (\mathbf{M} \mathbf{x}^{[K-2]} - \mathbf{m}) - \mathbf{m} \\ &= \mathbf{M}^2 \mathbf{x}^{[K-2]} - (\mathbf{M} + \mathbf{I}) \mathbf{m} \\ &= \mathbf{M}^2 (\mathbf{M} \mathbf{x}^{[K-3]} - \mathbf{m}) - (\mathbf{M} + \mathbf{I}) \mathbf{m} \\ &= \mathbf{M}^3 \mathbf{x}^{[K-3]} - (\mathbf{M}^2 + \mathbf{M} + \mathbf{I}) \mathbf{m}\end{aligned}$$

Matrix splitting: sketch of proof (1)

- Notations

$$\begin{aligned}\mathbf{x}^{[k]} &= \mathbf{A}^{-1} [\mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{q}] \\ &= \mathbf{A}^{-1} \mathbf{B} \mathbf{x}^{[k-1]} - \mathbf{A}^{-1} \mathbf{q} \\ &= \mathbf{M} \mathbf{x}^{[k-1]} - \mathbf{m}\end{aligned}$$

- Iterates

$$\begin{aligned}\mathbf{x}^{[K]} &= \mathbf{M} \mathbf{x}^{[K-1]} - \mathbf{m} \\ &= \mathbf{M} (\mathbf{M} \mathbf{x}^{[K-2]} - \mathbf{m}) - \mathbf{m} \\ &= \mathbf{M}^2 \mathbf{x}^{[K-2]} - (\mathbf{M} + \mathbf{I}) \mathbf{m} \\ &= \mathbf{M}^2 (\mathbf{M} \mathbf{x}^{[K-3]} - \mathbf{m}) - (\mathbf{M} + \mathbf{I}) \mathbf{m} \\ &= \mathbf{M}^3 \mathbf{x}^{[K-3]} - (\mathbf{M}^2 + \mathbf{M} + \mathbf{I}) \mathbf{m} \\ &= \mathbf{M}^K \mathbf{x}^{[0]} - (\mathbf{M}^{[K-1]} + \cdots + \mathbf{M}^2 + \mathbf{M} + \mathbf{I}) \mathbf{m} \\ &= \mathbf{M}^K \mathbf{x}^{[0]} - \left(\sum_{k=0}^{K-1} \mathbf{M}^k \right) \mathbf{m}\end{aligned}$$

Matrix splitting: sketch of proof (2)

- Convergence... in relation with eigenvalues of M

Matrix splitting: sketch of proof (2)

- Convergence... in relation with eigenvalues of M

- Increasing powers...

$$\begin{aligned} M &= P \Delta P^{-1} \\ M^2 &= P \Delta P^{-1} P \Delta P^{-1} = P \Delta^2 P^{-1} \\ M^K &= P \Delta^K P^{-1} \\ &\longrightarrow O \end{aligned}$$

Matrix splitting: sketch of proof (2)

- Convergence... in relation with eigenvalues of M

- Increasing powers...

$$\begin{aligned} M &= P \Delta P^{-1} \\ M^2 &= P \Delta P^{-1} P \Delta P^{-1} = P \Delta^2 P^{-1} \\ M^K &= P \Delta^K P^{-1} \\ &\longrightarrow O \end{aligned}$$

- ... and a well known serie

$$\begin{aligned} \sum_{k=0}^{K-1} M^k &= \sum_{k=0}^{K-1} (P \Delta P^{-1})^k \\ &= P \left(\sum_{k=0}^{K-1} \Delta^k \right) P^{-1} \end{aligned}$$

Matrix splitting: sketch of proof (2)

- Convergence... in relation with eigenvalues of M

- Increasing powers...

$$\begin{aligned} M &= P \Delta P^{-1} \\ M^2 &= P \Delta P^{-1} P \Delta P^{-1} = P \Delta^2 P^{-1} \\ M^K &= P \Delta^K P^{-1} \\ &\longrightarrow O \end{aligned}$$

- ... and a well known serie

$$\begin{aligned} \sum_{k=0}^{K-1} M^k &= \sum_{k=0}^{K-1} (P \Delta P^{-1})^k \\ &= P \left(\sum_{k=0}^{K-1} \Delta^k \right) P^{-1} \\ &\longrightarrow P (I - \Delta)^{-1} P^{-1} = (I - M)^{-1} \end{aligned}$$

- Convergence if: $\rho(M) < 1$, spectral radius of M smaller than 1

Matrix splitting: sketch of proof (3)

- Iterates

$$\boldsymbol{x}^{[K]} = \boldsymbol{M}^K \boldsymbol{x}^{[0]} - \left(\sum_{k=0}^{K-1} \boldsymbol{M}^k \right) \boldsymbol{m}$$

- Convergence if: $\rho(\boldsymbol{M}) < 1$, spectral radius of \boldsymbol{M} smaller than 1
- Limit for K tends to $+\infty$

$$\begin{aligned}\boldsymbol{x}^{\infty} &= \boldsymbol{M}^{\infty} \boldsymbol{x}^{[0]} - \sum_{k=0}^{\infty} \boldsymbol{M}^k \boldsymbol{m} \\ &= \boldsymbol{O} - (\boldsymbol{I} - \boldsymbol{M})^{-1} \boldsymbol{m} \\ &= -(\boldsymbol{I} - \boldsymbol{A}^{-1} \boldsymbol{B})^{-1} \boldsymbol{A}^{-1} \boldsymbol{q} \\ &= -(\boldsymbol{A} - \boldsymbol{B})^{-1} \boldsymbol{q} \\ &= -\boldsymbol{Q}^{-1} \boldsymbol{q}\end{aligned}$$

Matrix splitting: recap and examples

Matrix splitting recap

- Solve / compute

$$\begin{aligned} Qx &= -q \\ x &= -Q^{-1}q \end{aligned}$$

- Matrix splitting

$$Q = A - B$$

- Iterate

$$\begin{aligned} x^{[k]} &= A^{-1} [Bx^{[k-1]} - q] \\ Ax^{[k]} &= [Bx^{[k-1]} - q] \end{aligned}$$

- Computability:

- efficient inversion of A
- efficient solving of $Au = v$

- Convergence requires $\rho(A^{-1}B) < 1$

Matrix splitting: recap and examples

Matrix splitting examples

- Diagonal: $\mathbf{Q} = \mathbf{D} - \mathbf{R}$

- $\mathbf{x}^{[k]} = \mathbf{D}^{-1} [\mathbf{R} \mathbf{x}^{[k-1]} - \mathbf{q}]$

- Lower and Upper: $\mathbf{Q} = \mathbf{L}_+ - \mathbf{U}$

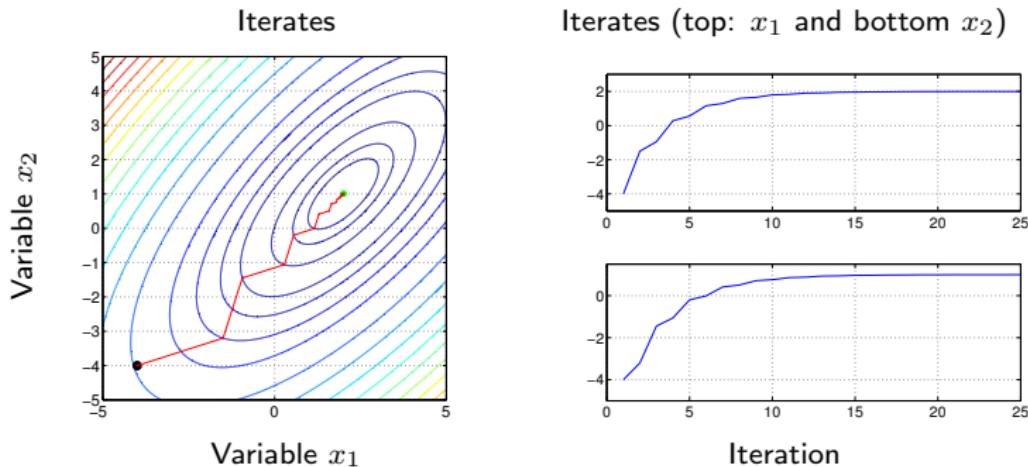
- $\mathbf{x}^{[k]} = \mathbf{L}_+^{-1} [\mathbf{U} \mathbf{x}^{[k-1]} - \mathbf{q}]$

- $\mathbf{L}_+ \mathbf{x}^{[k]} = [\mathbf{U} \mathbf{x}^{[k-1]} - \mathbf{q}]$

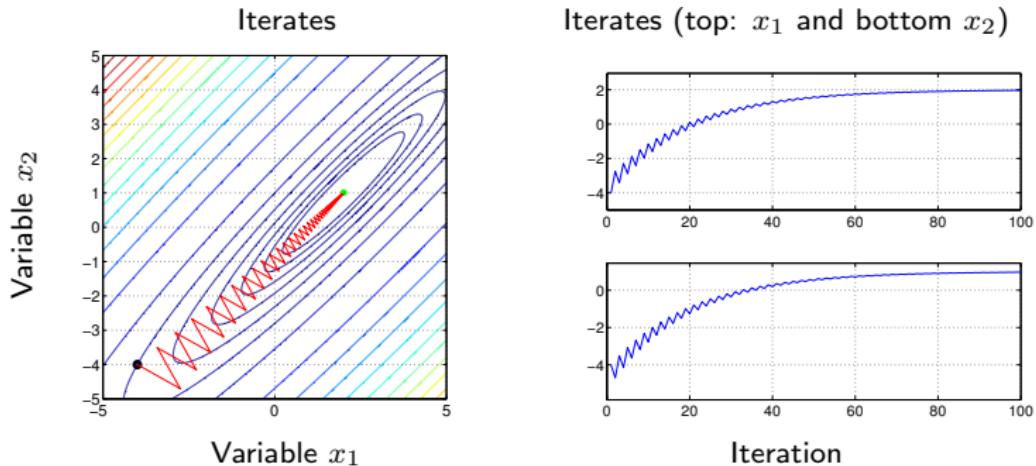
- Circulant: $\mathbf{Q} = \mathbf{C} - \mathbf{R}$

- $\mathbf{x}^{[k]} = \mathbf{C}^{-1} [\mathbf{R} \mathbf{x}^{[k-1]} - \mathbf{q}]$

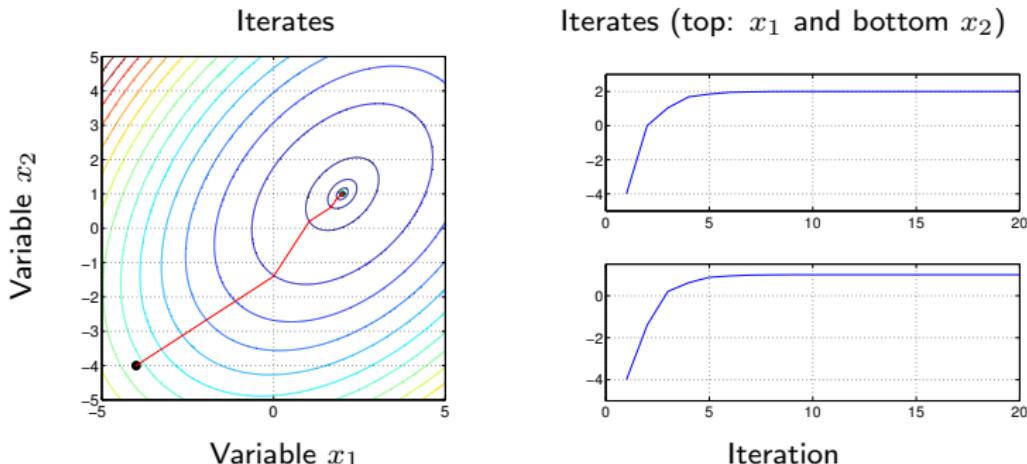
Matrix splitting: numerical results (1)



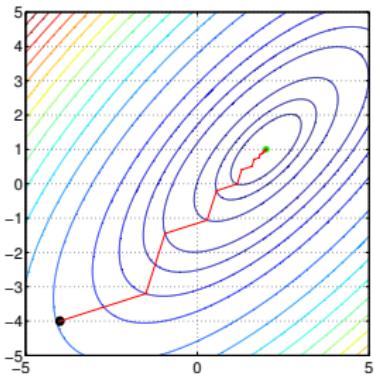
Matrix splitting: numerical results (2)



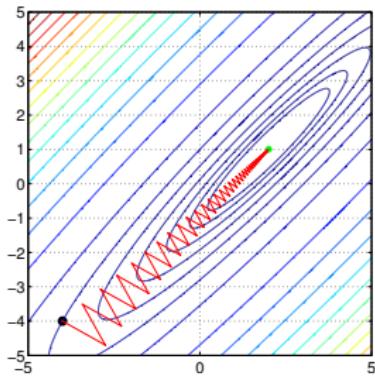
Matrix splitting: numerical results (3)



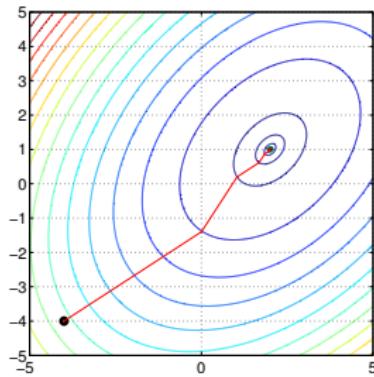
Matrix splitting: numerical results (all of them)



Coef. $\rho = 0.7$



Coef. $\rho = 0.95$



Coef. $\rho = 0.4$

Two families of **quadratic optimizers**

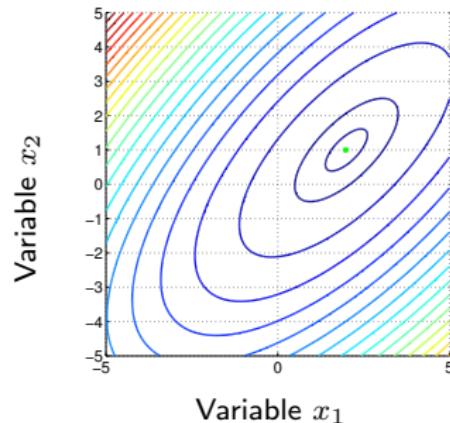
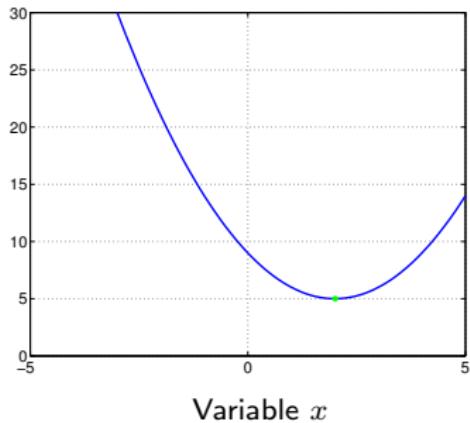
1 – **Component wise**

2 – **Gradient direction**

Quadratic criterion: illustration

① $\alpha^2 (x - \bar{x})^2 + \gamma$

② $\alpha_1^2 (x_1 - \bar{x}_1)^2 + \alpha_2^2 (x_2 - \bar{x}_2)^2 - 2\rho\alpha_1\alpha_2(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) + \gamma$



- Comment: convex, concave, saddle point, proper direction

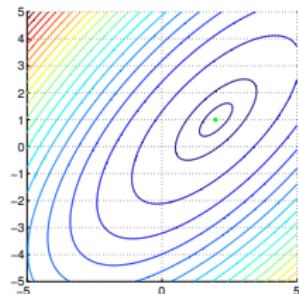
Minimisation, a first approach: component-wise

- Criterion

$$\mathcal{J}(\boldsymbol{x}) = \frac{1}{2} \boldsymbol{x}^t \mathbf{Q} \boldsymbol{x} + \mathbf{q}^t \boldsymbol{x} + q_0$$

- Iterative component-wise update

- For $k = 1, 2, \dots$
 - For $p = 1, 2, \dots P$
 - Update x_p by minimization of $\mathcal{J}(\boldsymbol{x})$ w.r.t. x_p given the other variables
 - End p
 - End k



- Properties

- Fixed point algorithm
- Convergence is proved

Component-wise minimisation (1)

- Criterion

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2} \mathbf{x}^t \mathbf{Q} \mathbf{x} + \mathbf{q}^t \mathbf{x} + q_0$$

- Two ingredients

- Select pixel p : vector $\mathbf{1}_p = [0, \dots, 0, 1, 0, \dots]^t$

$$x_p = \mathbf{1}_p^t \mathbf{x}$$

- Nullify pixel p : matrix $\mathbf{I} - \mathbf{1}_p \mathbf{1}_p^t$

$$\mathbf{x}_{/p} = (\mathbf{I} - \mathbf{1}_p \mathbf{1}_p^t) \mathbf{x}$$

- Rewrite current image...

$$\begin{aligned}\mathbf{x} &= (\mathbf{I} - \mathbf{1}_p \mathbf{1}_p^t) \mathbf{x} + x_p \mathbf{1}_p \\ &= \mathbf{x}_{/p} + x_p \mathbf{1}_p\end{aligned}$$

- ... and rewrite criterion

$$\tilde{\mathcal{J}}_p(x_p) = \mathcal{J} [\mathbf{x}_{/p} + x_p \mathbf{1}_p]$$

Component-wise minimisation (2)

- Consider the criterion as a function of pixel p

$$\begin{aligned}\tilde{\mathcal{J}}_p(x_p) &= \mathcal{J}[\mathbf{x}_{/p} + x_p \mathbf{1}_p] \\ &= \frac{1}{2}(\mathbf{x}_{/p} + x_p \mathbf{1}_p)^t \mathbf{Q} (\mathbf{x}_{/p} + x_p \mathbf{1}_p) + \mathbf{q}^t (\mathbf{x}_{/p} + x_p \mathbf{1}_p) + q_0 \\ &\quad \vdots \\ &= \frac{1}{2} \mathbf{1}_p^t \mathbf{Q} \mathbf{1}_p x_p^2 + (\mathbf{Q} \mathbf{x}_{/p} + \mathbf{q})^t \mathbf{1}_p x_p + \dots\end{aligned}$$

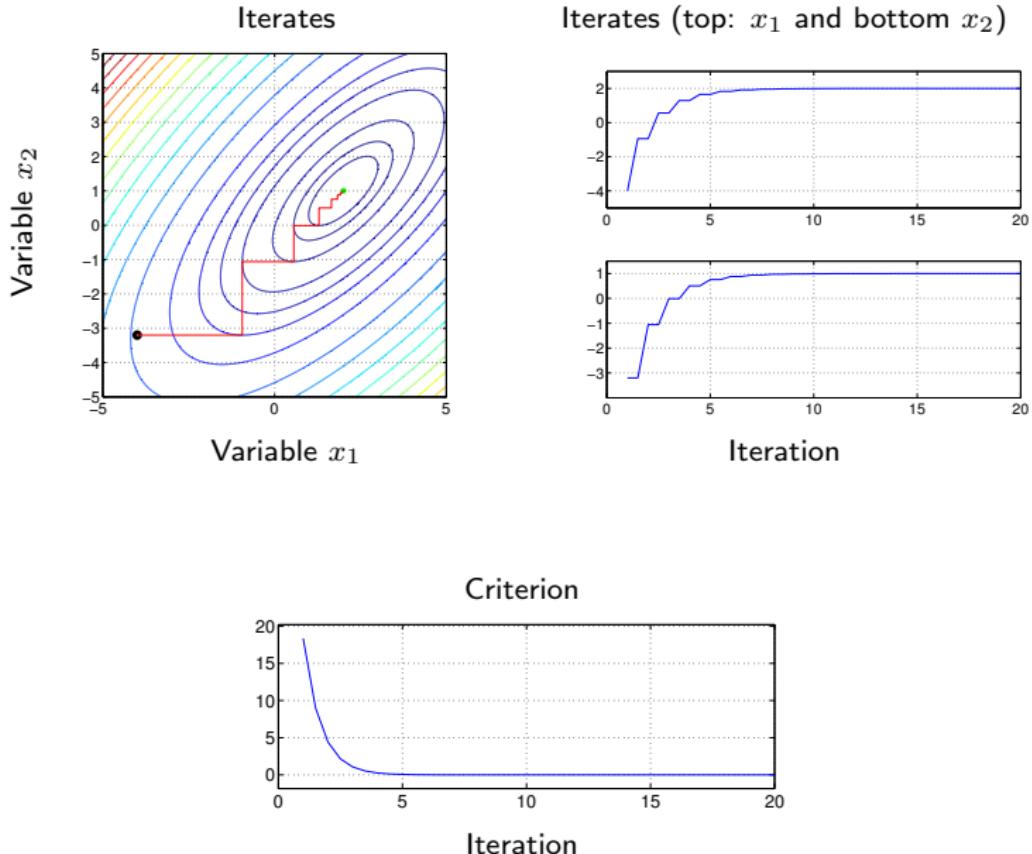
- ... and minimize: update pixel p

$$x_p^{\text{opt}} = -\frac{(\mathbf{Q} \mathbf{x}_{/p} + \mathbf{q})^t \mathbf{1}_p}{\mathbf{1}_p^t \mathbf{Q} \mathbf{1}_p}$$

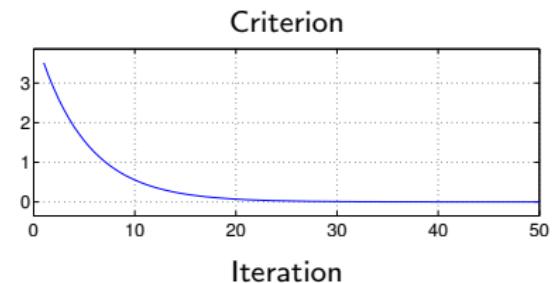
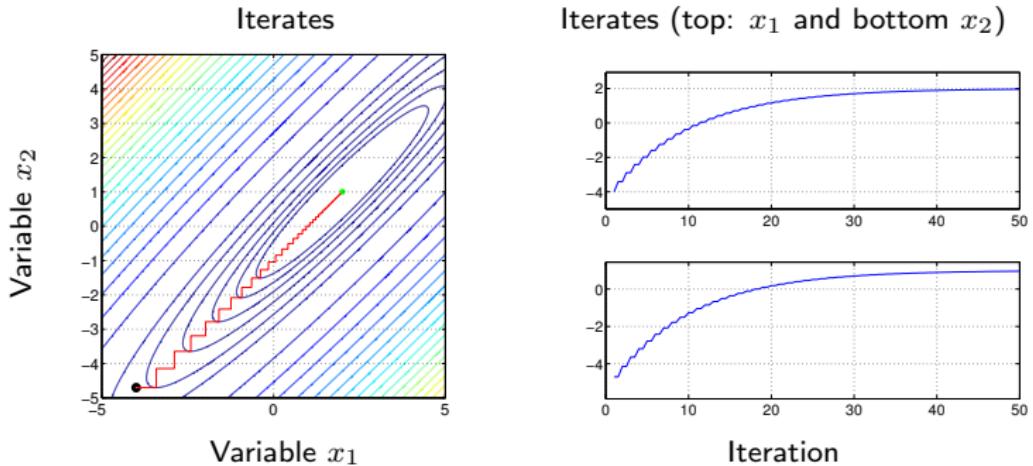
- Computation load

- Practicability of computing $\mathbf{1}_p^t \mathbf{Q} \mathbf{1}_p$ (think about side-effects)
- Possible efficient update of $\phi_p = (\mathbf{Q} \mathbf{x}_{/p} + \mathbf{q})^t \mathbf{1}_p$ itself

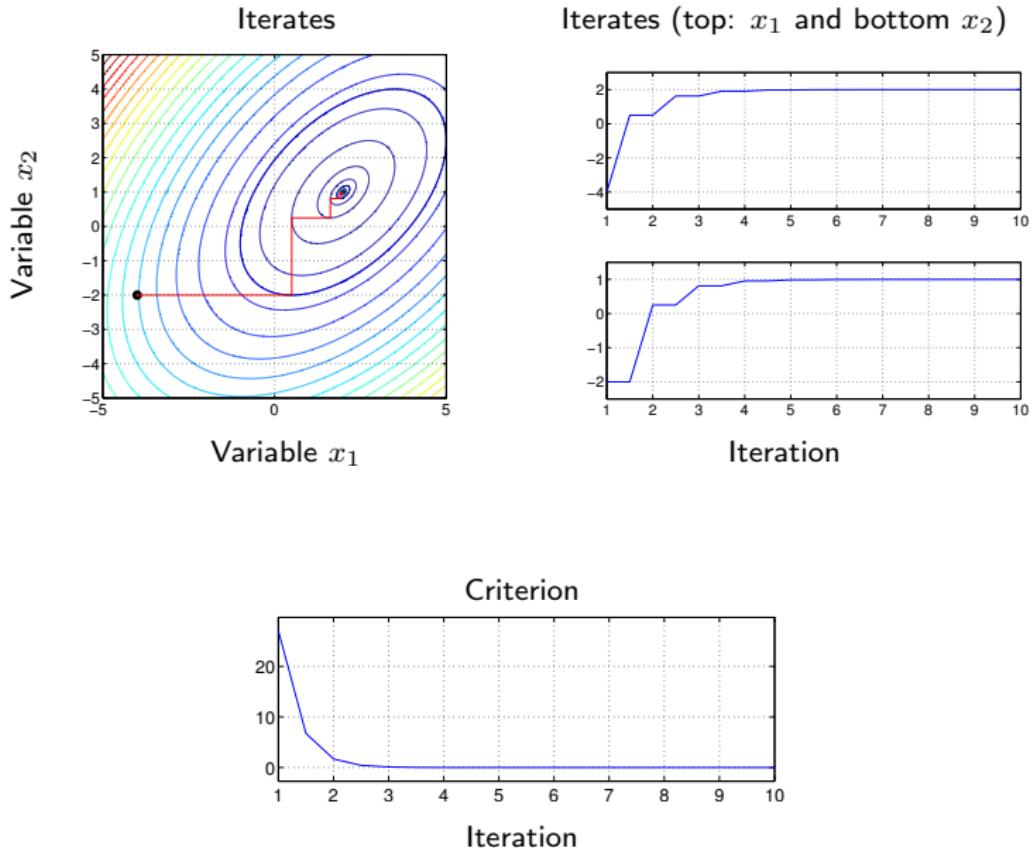
Component-wise: numerical results (1)



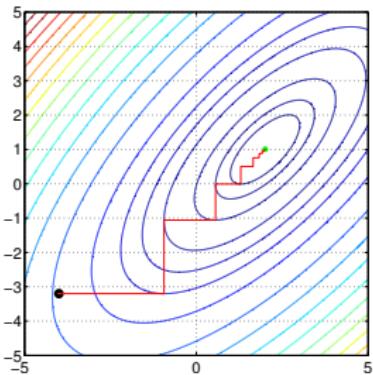
Component-wise: numerical results (2)



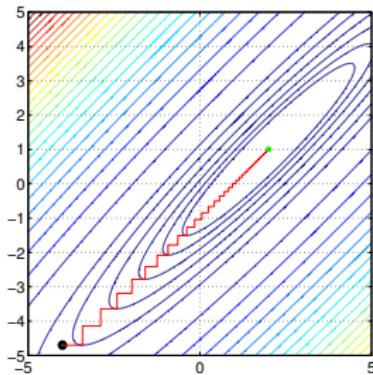
Component-wise: numerical results (3)



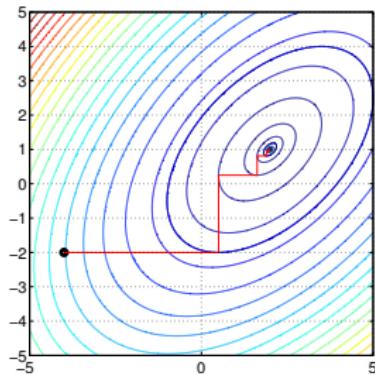
Component-wise: numerical results (all of them)



Coef. $\rho = 0.7$



Coef. $\rho = 0.95$



Coef. $\rho = 0.4$

Extensions: multipixel and separability

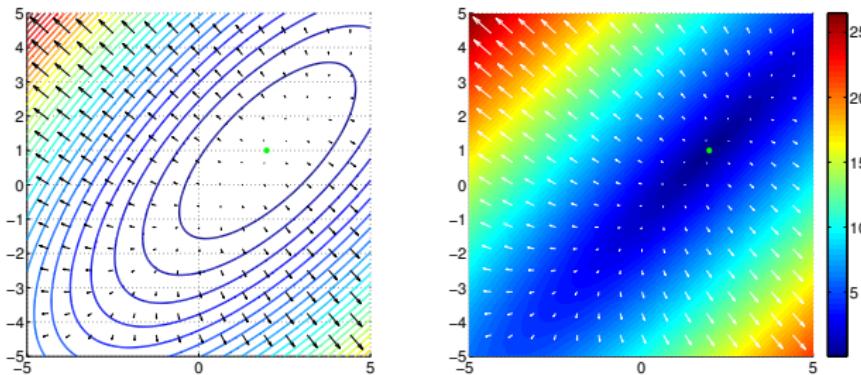
Two families of **quadratic optimizers**

1 – **Component wise**

2 – **Gradient direction**

Quadratic criterion: generalities

- $\alpha_1^2 (x_1 - \bar{x}_1)^2 + \alpha_2^2 (x_2 - \bar{x}_2)^2 - 2\rho\alpha_1\alpha_2(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) + \gamma$



Optimality over \mathbb{R}^N (no constraint)

- Null gradient

$$\mathbf{0} = \left. \frac{\partial \mathcal{J}}{\partial \mathbf{x}} \right|_{\bar{\mathbf{x}}} = \mathbf{g}(\bar{\mathbf{x}})$$

- Positive Hessian

$$\mathbf{Q} = \frac{\partial^2 \mathcal{J}}{\partial \mathbf{x}^2} > 0$$

Iterative algorithms, fixed point,...

$$\bar{\boldsymbol{x}} = \arg \min_{\boldsymbol{x}} \mathcal{J}(\boldsymbol{x})$$

Iterative update

- Initialisation $\boldsymbol{x}^{[0]}$
- Iteration $k = 0, 1, 2, \dots$

$$\boldsymbol{x}^{[k+1]} = \boldsymbol{x}^{[k]} + \tau^{[k]} \boldsymbol{\delta}^{[k]}$$

- Direction $\boldsymbol{\delta} \in \mathbb{R}^N$
- Step length $\tau \in \mathbb{R}_+$

$$\boxed{\boldsymbol{x}^{[k]} \xrightarrow[k \rightarrow \infty]{} \bar{\boldsymbol{x}}}$$

- Stopping rule, e.g. $\|\boldsymbol{g}(\boldsymbol{x}^{[k]})\| < \varepsilon$

- Direction δ
 - “Optimal”: opposite of the gradient
 - Newton, inverse Hessian
 - Preconditioned
 - Corrected directions:
 - bisector, Vignes, Polak-Ribière,...
 - conjugate direction,...
 - ...
- Step length τ
 - “Optimal”
 - Over-relaxed / under-relaxed
 - Armijo, Goldstein, Wolfe
 - ...
 - and also fixed step
- ... optimal, yes,... and no...

Gradient with optimal step length

Strategy: optimal direction \oplus optimal step length

- Iteration $\boldsymbol{x}^{[k+1]} = \boldsymbol{x}^{[k]} + \tau^{[k]} \boldsymbol{\delta}^{[k]}$

- Direction $\boldsymbol{\delta} \in \mathbb{R}^N$

$$\boldsymbol{\delta}^{[k]} = -\mathbf{g}(\boldsymbol{x}^{[k]})$$

- Step length $\tau \in \mathbb{R}_+$

$$\mathcal{J}_{\boldsymbol{\delta}}(\tau) = \mathcal{J}(\boldsymbol{x}^{[k]} + \tau \boldsymbol{\delta}^{[k]})$$

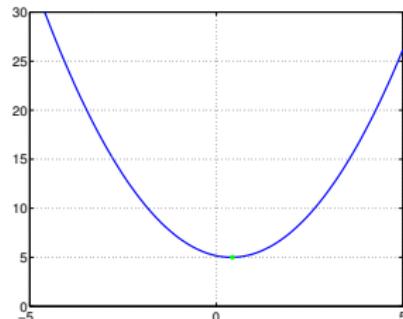
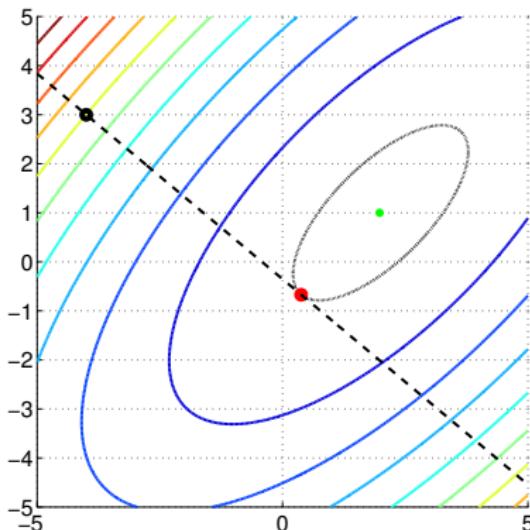
- A second order polynom

$$\mathcal{J}_{\boldsymbol{\delta}}(\tau) = \dots \tau^2 + \dots \tau + \dots$$

- Optimal step length

$$\tau^{[k]} = \frac{\mathbf{g}(\boldsymbol{x}^{[k]})^t \mathbf{g}(\boldsymbol{x}^{[k]})}{\mathbf{g}(\boldsymbol{x}^{[k]})^t \mathbf{Q} \mathbf{g}(\boldsymbol{x}^{[k]})}$$

Illustration



- Two readings:
 - optimisation along the given direction, “line-search”
 - constrained optimisation
- Remark: orthogonality of successive directions (see end of exercise)

Sketch of convergence proof (0)

A reminder for notations

$$\begin{aligned}\mathcal{J}(\mathbf{x}) &= \frac{1}{2} \mathbf{x}^t \mathbf{Q} \mathbf{x} + \mathbf{q}^t \mathbf{x} + q_0 \\ \mathbf{g}(\mathbf{x}) &= \mathbf{Q} \mathbf{x} + \mathbf{q}\end{aligned}$$

$$\bar{\mathbf{x}} = \arg \min_{\mathbf{x}} \mathcal{J}(\mathbf{x}) = -\mathbf{Q}^{-1} \mathbf{q}$$

Some preliminary results

$$\mathcal{J}(\bar{\mathbf{x}}) = -\frac{1}{2} \mathbf{q}^t \mathbf{Q}^{-1} \mathbf{q} + q_0$$

$$\mathcal{J}(\mathbf{x}) - \mathcal{J}(\bar{\mathbf{x}}) = \frac{1}{2} \mathbf{g}(\mathbf{x})^t \mathbf{Q}^{-1} \mathbf{g}(\mathbf{x})$$

$$\mathcal{J}(\mathbf{x}_0 + \mathbf{h}) = \mathcal{J}(\mathbf{x}_0) + \mathbf{g}(\mathbf{x}_0)^t \mathbf{h} + \frac{1}{2} \mathbf{h}^t \mathbf{Q} \mathbf{h}$$

Sketch of convergence proof (1)

- Notation convenience: $\mathbf{x}^{[k]} = \mathbf{x}$, $\mathbf{x}^{[k+1]} = \mathbf{x}'$, $\tau^{[k]} = \tau$, $\mathbf{g}(\mathbf{x}^{[k]}) = \mathbf{g}$
- Criterion increase / decrease

$$\begin{aligned}\mathcal{J}(\mathbf{x}') - \mathcal{J}(\mathbf{x}) &= \mathcal{J}(\mathbf{x} + \tau \boldsymbol{\delta}) - \mathcal{J}(\mathbf{x}) \\ &= \mathbf{g}^t [\tau \boldsymbol{\delta}] + \frac{1}{2} [\tau \boldsymbol{\delta}]^t \mathbf{Q} [\tau \boldsymbol{\delta}] \\ &= \dots \\ &= -\frac{1}{2} \frac{[\mathbf{g}^t \mathbf{g}]^2}{\mathbf{g}^t \mathbf{Q} \mathbf{g}}\end{aligned}$$

since $\boldsymbol{\delta} = -\mathbf{g}$ and $\tau = \frac{\mathbf{g}^t \mathbf{g}}{\mathbf{g}^t \mathbf{Q} \mathbf{g}}$

- Comment: it decreases, it reduces...

Sketch of convergence proof (2)

Distance to minimiser

$$\begin{aligned}\mathcal{J}(\mathbf{x}') - \mathcal{J}(\bar{\mathbf{x}}) &= \mathcal{J}(\mathbf{x}) - \mathcal{J}(\bar{\mathbf{x}}) - \frac{1}{2} \frac{[\mathbf{g}^t \mathbf{g}]^2}{\mathbf{g}^t \mathbf{Q} \mathbf{g}} \\ &= \mathcal{J}(\mathbf{x}) - \mathcal{J}(\bar{\mathbf{x}}) - \frac{1}{2} \frac{[\mathbf{g}^t \mathbf{g}]^2}{\mathbf{g}^t \mathbf{Q} \mathbf{g}} \times \frac{\mathcal{J}(\mathbf{x}) - \mathcal{J}(\bar{\mathbf{x}})}{\mathbf{g}^t \mathbf{Q}^{-1} \mathbf{g} / 2} \\ &= [\mathcal{J}(\mathbf{x}) - \mathcal{J}(\bar{\mathbf{x}})] \left[1 - \frac{[\mathbf{g}^t \mathbf{g}]^2}{[\mathbf{g}^t \mathbf{Q} \mathbf{g}] [\mathbf{g}^t \mathbf{Q}^{-1} \mathbf{g}]} \right] \\ &= \dots \\ &\leq [\mathcal{J}(\mathbf{x}) - \mathcal{J}(\bar{\mathbf{x}})] \left(\frac{M - m}{M + m} \right)^2\end{aligned}$$

... so "it converges"

- M and m : maximal and minimal eigenvalue of \mathbf{Q} ... and comment
- Kantorovich inequality (see next slide)

Kantorovich inequality

Result

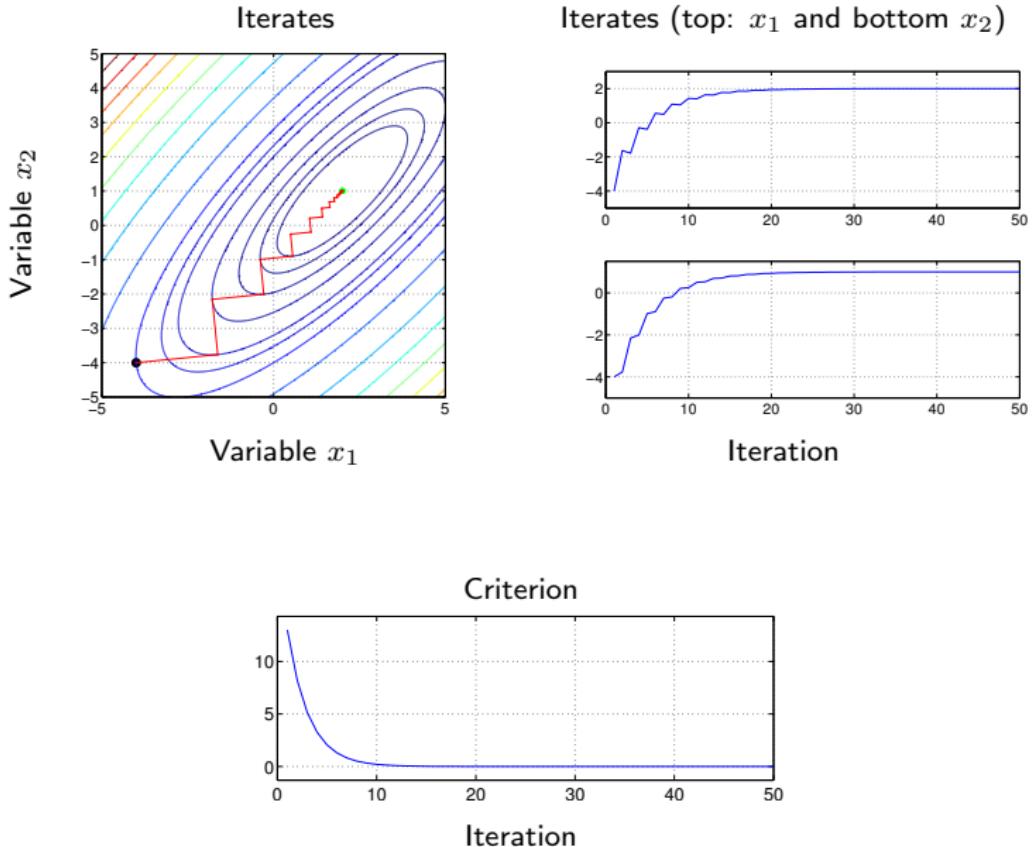
$$[u^t Q u] [u^t Q^{-1} u] \leq \frac{1}{4} \left(\sqrt{\frac{m}{M}} + \sqrt{\frac{M}{m}} \right)^2 \|u\|^4$$

- Q symmetric and positive-definite
- M and m : maximal and minimal eigenvalue of Q

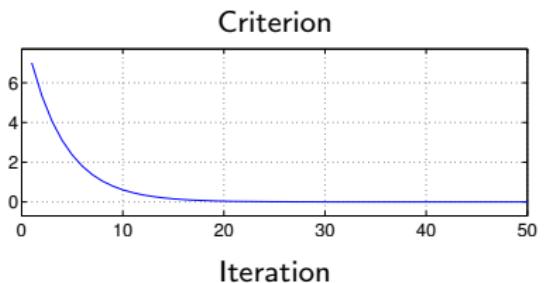
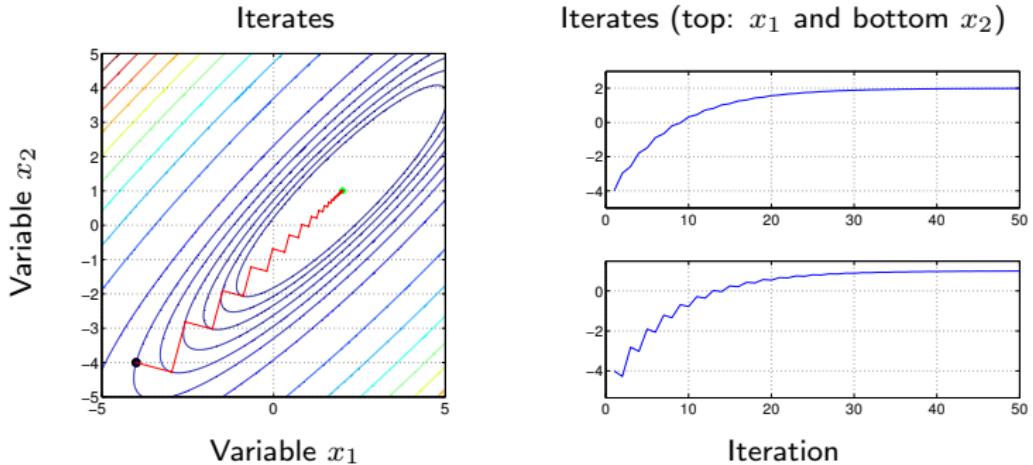
Short sketch of proof

- Quite long and complex
- Case $\|u\| = 1$
- Diagonalise Q : $Q = P^t \Lambda P$ et $Q^{-1} = P^t \Lambda^{-1} P$
- Convex combination of the eigenvalues and their inverse
- Convexity of $t \mapsto 1/t$

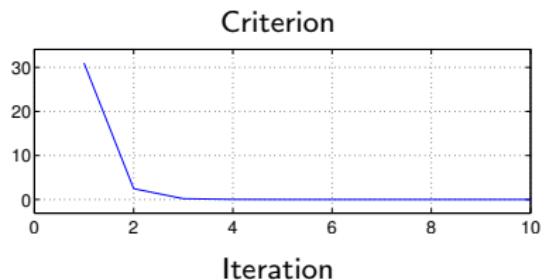
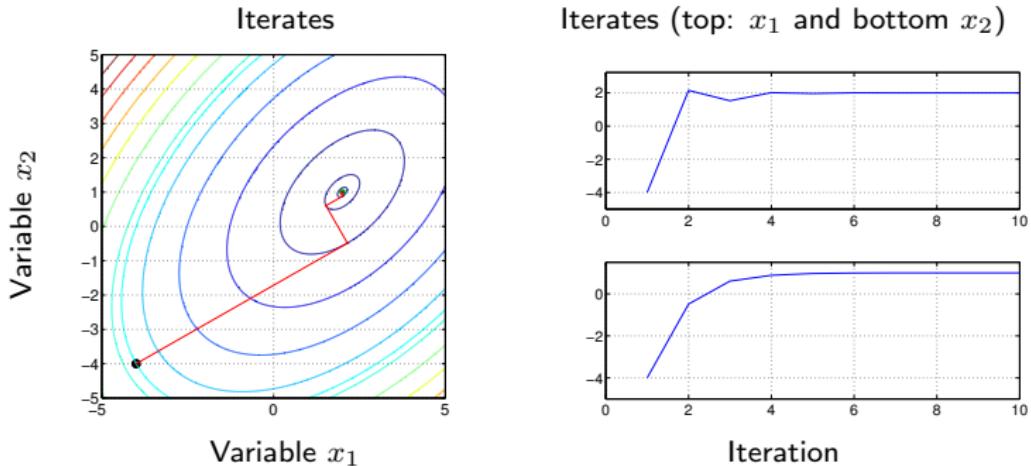
Optimal gradient: numerical results (1)



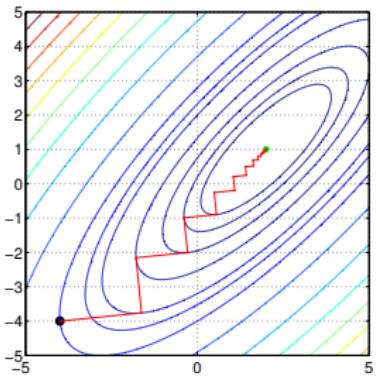
Optimal gradient: numerical results (2)



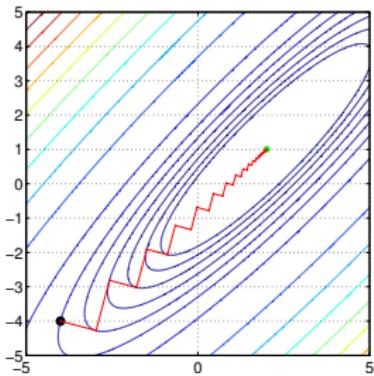
Optimal gradient: numerical results (3)



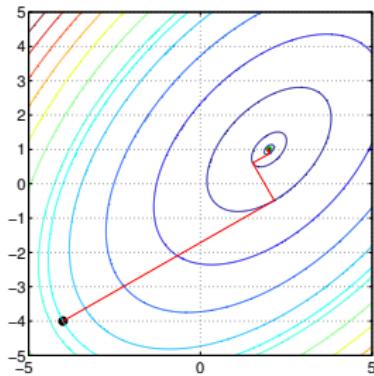
Optimal gradient: numerical results (all of them)



Coef. $\rho = 0.7$



Coef. $\rho = 0.95$



Coef. $\rho = 0.4$

Preconditioned gradient: short digest...

Strategy: modified direction \oplus optimal step length

- Iteration $\boldsymbol{x}^{[k+1]} = \boldsymbol{x}^{[k]} + \tau^{[k]} \boldsymbol{\delta}^{[k]}$

- Direction $\boldsymbol{\delta} \in \mathbb{R}^N$

$$\boldsymbol{\delta}^{[k]} = -\mathbf{P}\mathbf{g}(\boldsymbol{x}^{[k]})$$

- Step length $\tau \in \mathbb{R}_+$

$$\mathcal{J}_{\boldsymbol{\delta}}(\tau) = \mathcal{J}(\boldsymbol{x}^{[k]} + \tau \boldsymbol{\delta}^{[k]})$$

- A second order polynom

$$\mathcal{J}_{\boldsymbol{\delta}}(\tau) = \dots \tau^2 + \dots \tau + \dots$$

- Optimal step length

$$\tau^{[k]} = \frac{\mathbf{g}(\boldsymbol{x}^{[k]})^\text{t} \mathbf{P} \mathbf{g}(\boldsymbol{x}^{[k]})}{\mathbf{g}(\boldsymbol{x}^{[k]})^\text{t} \mathbf{P} \mathbf{Q} \mathbf{P} \mathbf{g}(\boldsymbol{x}^{[k]})}$$

Preconditioned gradient: two special cases . . .

- Non-preconditioned: $P = I \rightsquigarrow$ standard gradient algorithm

- Direction $\delta \in \mathbb{R}^N$

$$\delta^{[k]} = -\mathbf{g}(\mathbf{x}^{[k]})$$

- Step length $\tau \in \mathbb{R}_+$

$$\tau^{[k]} = \frac{\mathbf{g}(\mathbf{x}^{[k]})^\text{t} \mathbf{g}(\mathbf{x}^{[k]})}{\mathbf{g}(\mathbf{x}^{[k]})^\text{t} \mathbf{Q} \mathbf{g}(\mathbf{x}^{[k]})}$$

- Perfect-preconditioner: $P = Q^{-1} \rightsquigarrow$ one step optimal

- Direction $\delta \in \mathbb{R}^N$

$$\delta^{[k]} = -\mathbf{Q}^{-1} \mathbf{g}(\mathbf{x}^{[k]})$$

- Step length $\tau \in \mathbb{R}_+$

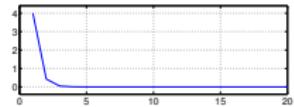
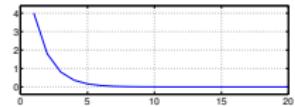
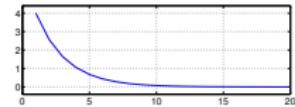
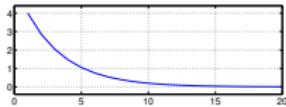
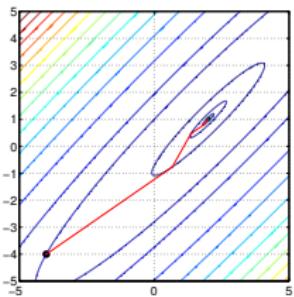
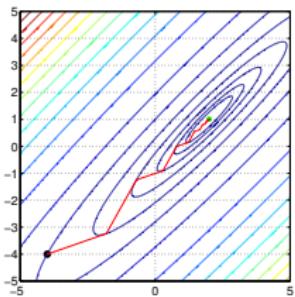
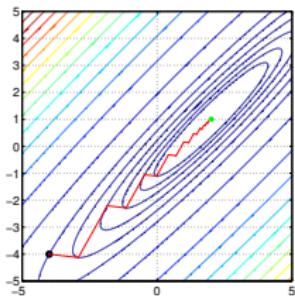
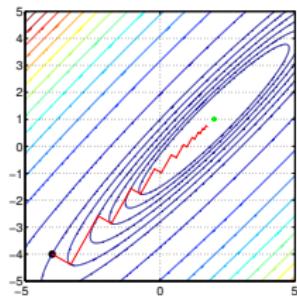
$$\tau^{[k]} = \frac{\mathbf{g}(\mathbf{x}^{[k]})^\text{t} P \mathbf{g}(\mathbf{x}^{[k]})}{\mathbf{g}(\mathbf{x}^{[k]})^\text{t} P Q P \mathbf{g}(\mathbf{x}^{[k]})} = 1$$

- First iterate

$$\begin{aligned}\mathbf{x}^{[1]} &= \mathbf{x}^{[0]} - 1 \cdot \mathbf{Q}^{-1} \mathbf{g}(\mathbf{x}^{[0]}) \\ &= \mathbf{x}^{[0]} - 1 \cdot \mathbf{Q}^{-1} (\mathbf{Q} \mathbf{x}^{[0]} + \mathbf{q}) = -\mathbf{Q}^{-1} \mathbf{q} !\end{aligned}$$

Preconditionned gradient: numerical results:

- Better and better approximation of the inverse Hessian



- Image restoration, deconvolution and other inverse problems
- Three types of regularised inversion...
 - Quadratic penalties
 - Convex non-quadratic penalties
 - Constraints: positivity and support
- ... based on unconstrained quadratic optimisation
- A basic component: Quadratic criterion and Gaussian model
 - Circulant approximation and computations based on FFT
 - Numerical linear system solvers
 - Matrix splitting
 - Numerical quadratic optimizers
 - Component-wise
 - Gradient methods

Various solutions...



True



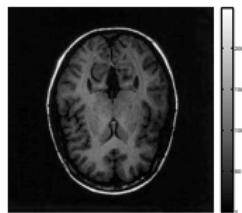
Observation



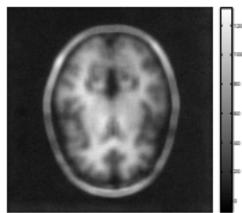
Quadratic penalty



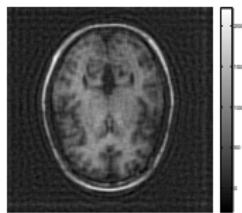
Huber penalty



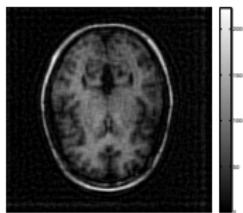
True



Observation



Quadratic penalty



Constrained