ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Model-based reconstruction methods

Computational imaging
2024-25

Elena Loli Piccolomini

Dipartimento di Informatica - Scienza e Ingegneria (DISI)

Direct regularization

Direct regularization methods compute the solution of an ill posed inverse problem (imaging) by imposing

Some constraints on the solution.

The so called *model-based* approach mathematically models the problem to be solved as a **minimization**

Problem with two acting functions:

- The term $\|Ax-(y+e)\|_2^2 = \|Ax-y^\delta\|_2^2$, representing the data fitting
- The regularization term R(x) that incrporates a priori infromation on the solution.

Direct regularization

The minimization can be expressed as a constrained minimization:

$$min R(x) \text{ such that } \|Ax - y^{\delta}\|_2^2 = \epsilon$$

Or

$$min \|Ax - y^{\delta}\|_2^2 \text{ such that } R(x) = \sigma$$

Or as an equivalent unconstrained minimization:

$$min \|Ax - (y + e)\|_2^2 + \lambda R(x)$$

where λ is the regularization parameter representing the trade off between the fit-to-data and the regularization terms.

Some popular regularizers

The p-norm regularization

$$R(x) = \|Lx\|_p^p, \quad 0 < p \leq 2$$

Where the p-orm is defined as:

$$\|x\|_p^p = (\sum_{i=1}^n x_i^p)^{1/p}$$

The most popular settings are:

- Concerning the operator L: $L = I$ or $L = D$ where D is the discrete gradient (or derivative) operator
- Concerning the exponent p: p=1, p=2

The *gradient* of an image

Given a differentiable function $f : R^2 \to R$, we know that the gradient of f $\nabla f : R^2 \to R^2$

is a function Constituted by the partial derivatives of f:

$$\nabla f(x) = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2})^T$$

We remark that if all the partial derivatives of a function $f : R^n \to R$ exist and are are conitnuous

in a point x_0 then f is said differentiable in x_0.

However, the image is not a function, but a matrix. How can we «define» the gradient of an image?

WE cosnider the discretization of the gradient function and apply it to the pixels of the image.

The *gradient* of an image

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

The image gradient

image

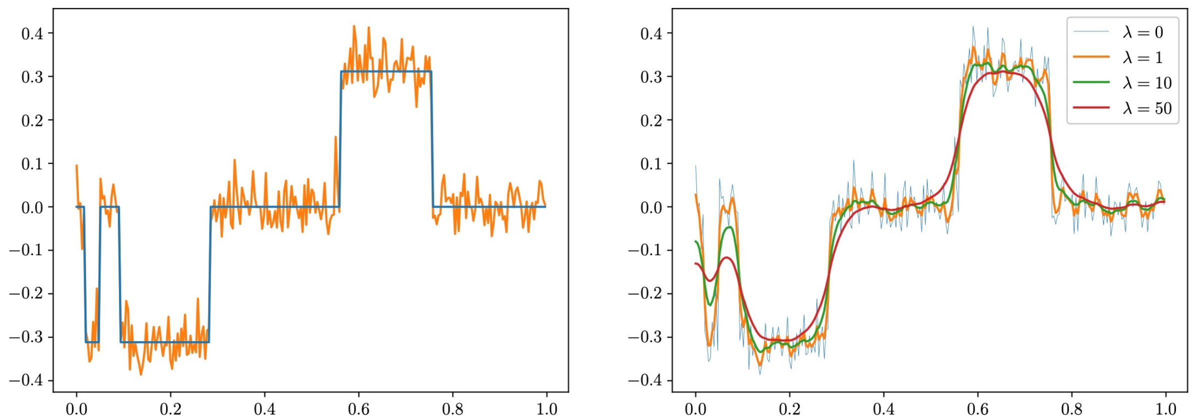
gradient

Tikhonov regularization method: p=2

Tikhonov corresponds to the choice p=2.

$$min\|Ax - y^\delta\|_2^2 + \lambda\|Lx\|_2^2$$

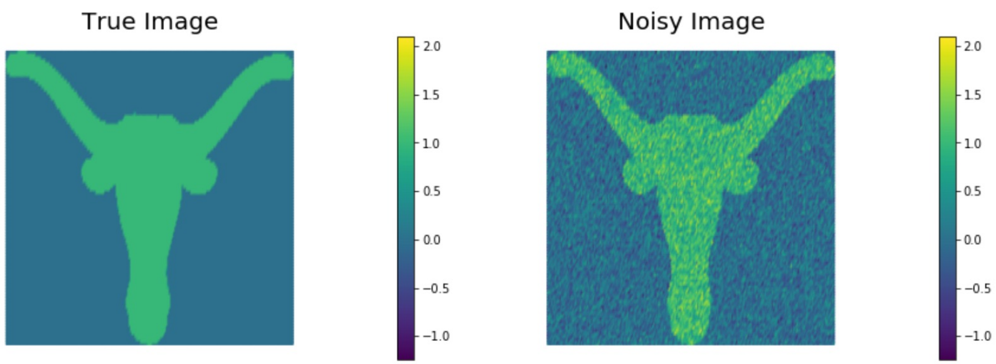
- Tikhonov regularization imposes smoothness on the solution (in the following examples L=I).



(a) Noisy signal (b) Tikhonov denoising

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Tikhonov regularization



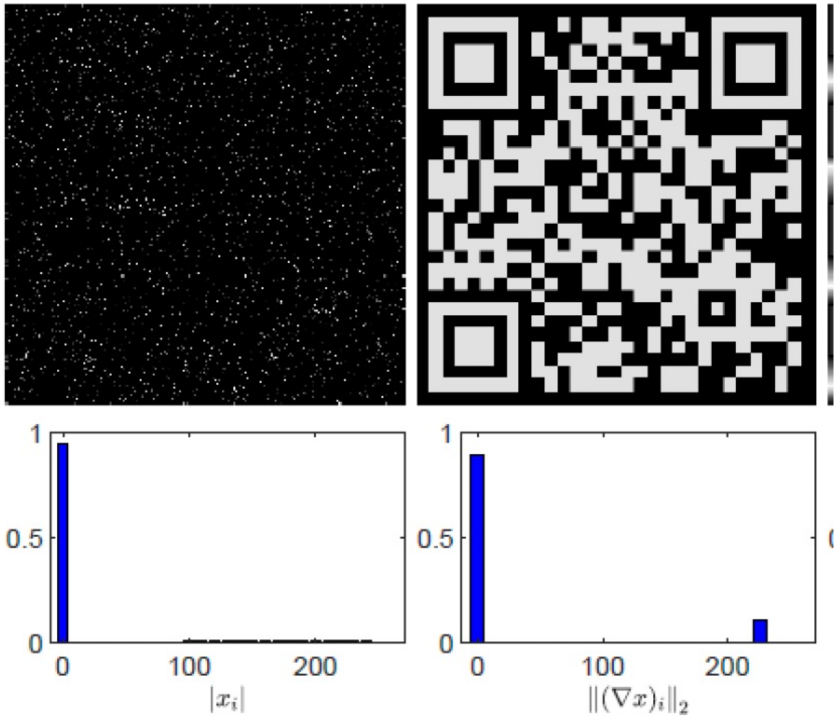
Tikhonov regularization

$\lambda = 10^{-2}$ $\lambda = 10^{-3}$ $\lambda = 10^{-4}$

Sparsifing regularizers and compressed sensing

A new challenge in imaging is to reconstruct an image which is sparse in some domain.

Sparse in the image domain



Sparse in the gradient domain

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Compressed sensing

Compressed Sensing is a recent methodology of signal and image acquisition and reconstruction
[Candes et al, IEEE transaction o Infromation theory, vol. 52, 2006,
Donoho, IEEE Trans. On Inf. Thoery, vol 52, 2006]

When the signal (or image) is sparse in a certain domain.

Sparse representation means that a signal can be represented with a few significant non-zero components.

In this case the problem is formulated as:

$$min\|Lx\|_1 \text{ so that } Ax = y^\delta$$

Or in a relaxed fromulation:

$$min\|Lx\|_1 \text{ so that } \|Ax - y^\delta\|_2^2 = \epsilon$$

Where epsilon is an estimate of the noise norm.

Total Variation regularization: L=gradient

If we consider L as the gradient of the image, the equivalent Lasso unconstrained problem is :

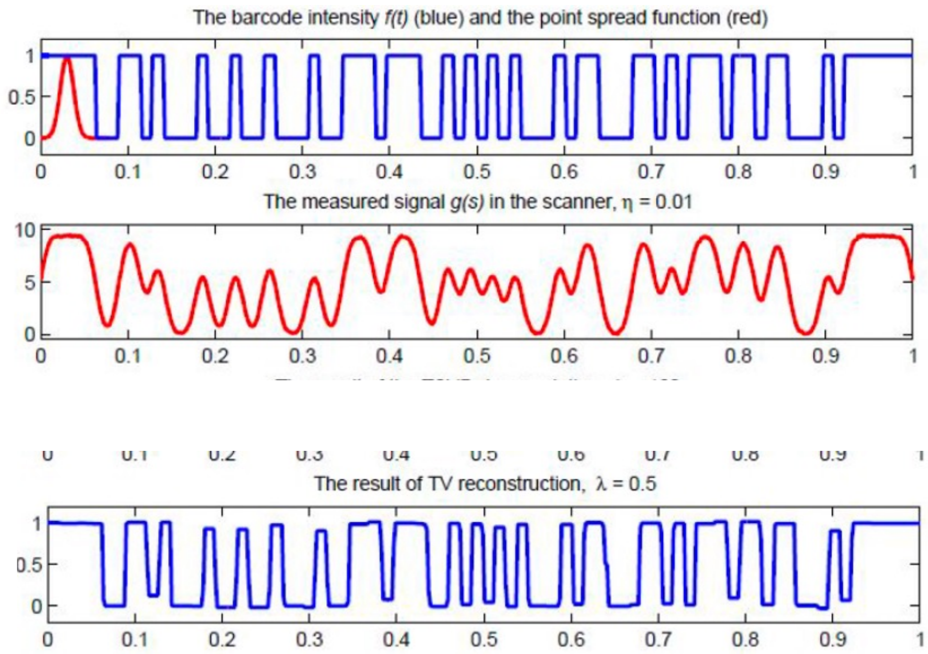
$$min\|Ax-(y+e)\|_2^2+\lambda\|\nabla x\|_1$$

The regularization term is called Total Variation (TV) of x.

- .Isotropic TV: $TV(x)=\sum_{i=1}^{n}\sqrt{(D_hx)_i^2+(D_vx)_i^2}$

Where D_hx Is the matrix of the discrete horizontal partial derivative of x and D_vx is the matrix of the discrete vertical partial derivative of x.

Total Variation regularization,p=1

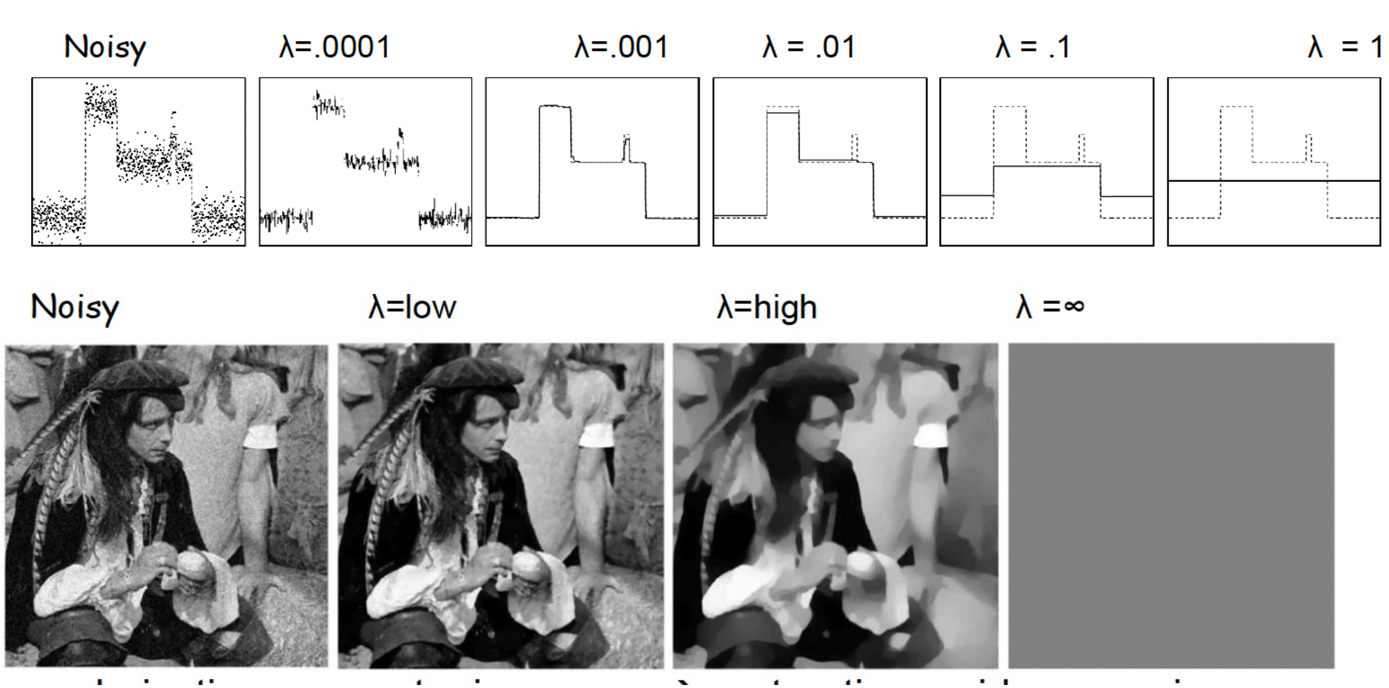


The Total Variation is suitable for signals and images sparse in the gradient domain.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

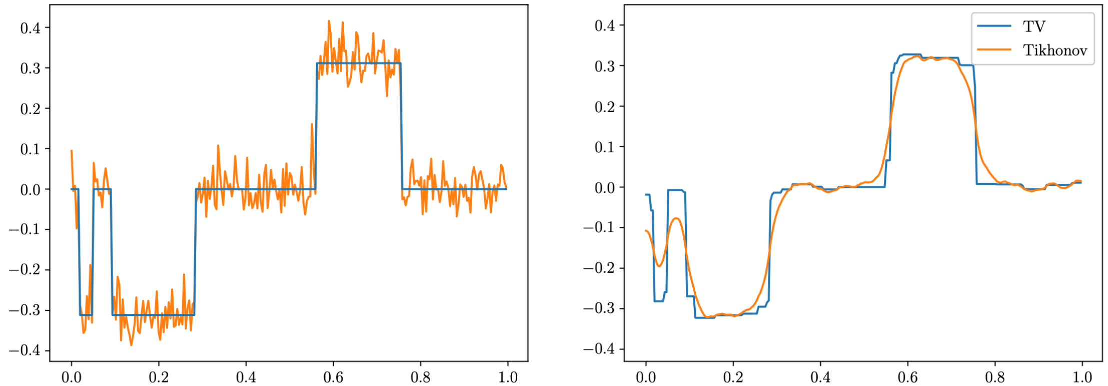
Total Variation regularization, p=1

Effects of the regularization parameter in TV regularization:



TV is a regularizer suitable for piecewise constant signals and images.

TV vs Tikhonov regularization



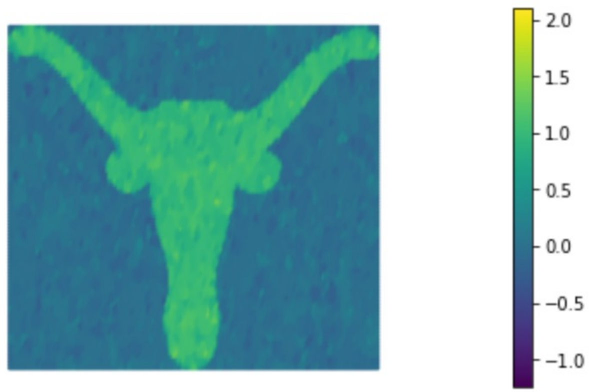
(c) Noisy signal (d) Denoised

It turns out that TV(x) really is a powerful method, but numerical minimization is more difficult than in the case of Tikhonov regularization; this is because the function to be minimized is no more quadratic (and actually noteven differentiable).

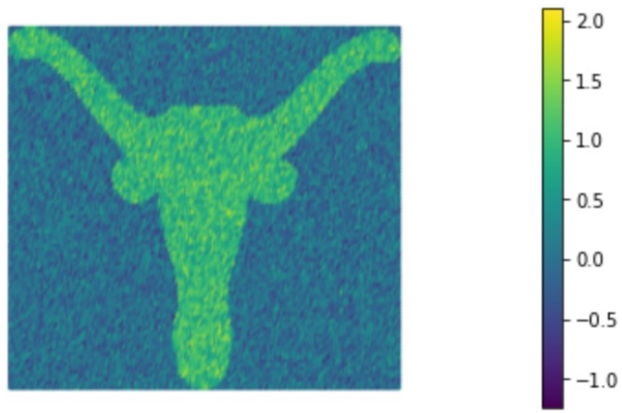
TV regularization

TV regularization

$\lambda = 10^{-2}$



$\lambda = 10^{-3}$



$\lambda = 10^{-4}$

Case p=0

When p=0 we obtain a semi-norm (not all the properties of a norm are satisfied)

$|z|_0 = p$ If p is the number of non-zero elements in z.

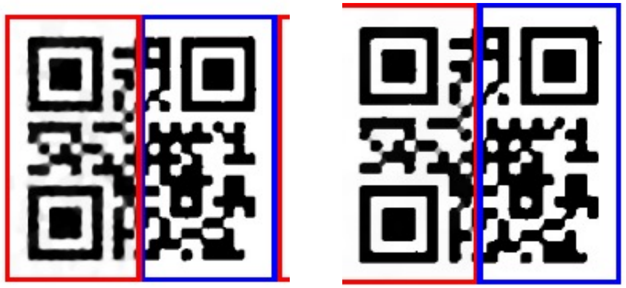
Constrained formulation with L=I: $min\|x\|_0 \text{ such that } Ax = y^\delta$

Unconstrained formulation with L=D:

$$TV^0(x) = \sum_{i=1}^{N} |\sqrt{(D_h x)_i^2 + (D_v x)_i^2}|_0$$

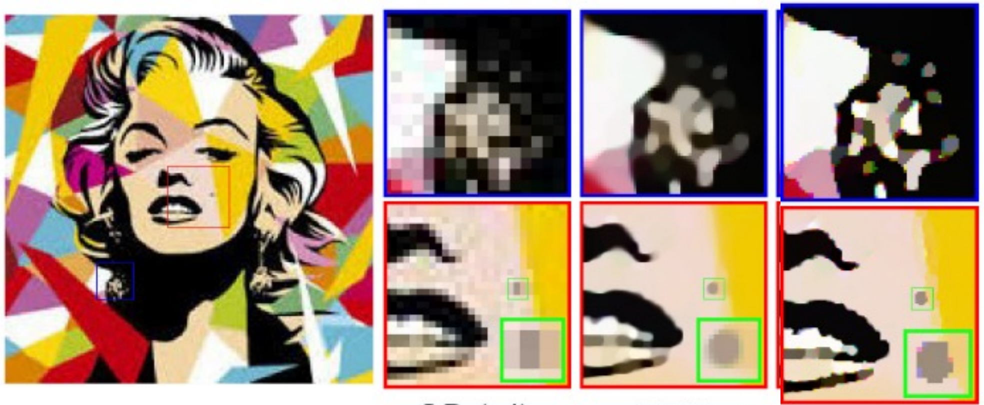
Total Variation regularization, p=0

Very sparse signal in the gradient domain, Super Resolution



TV   TV^0

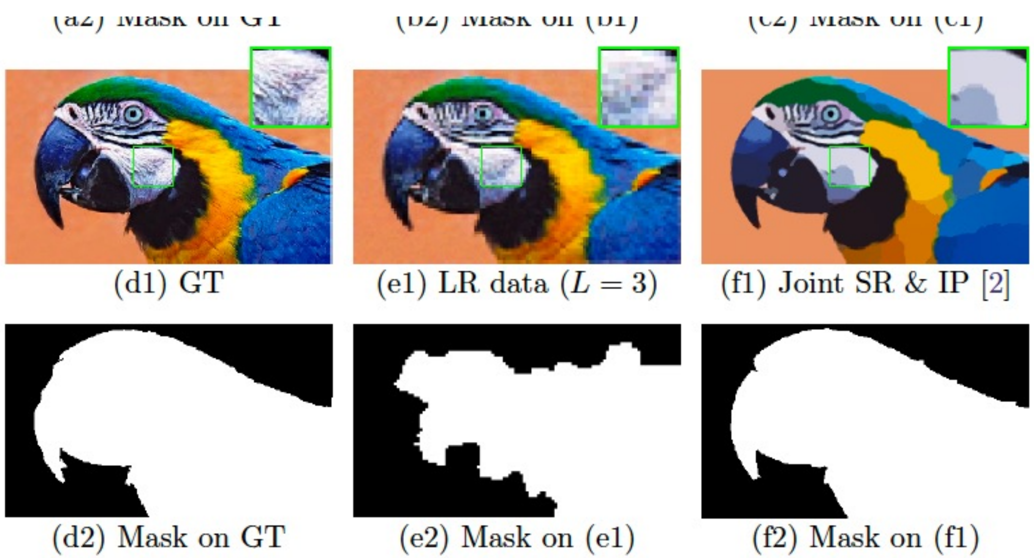
LR (x4)   I-TV   A-TV^0



LR (x4)   I-TV   I-TV^0

Total Variation regularization, p=0

Super Resolution and mask decetion with TV_0



Combined regularization

$$min\|Ax - y^\delta\|_2^2 + \lambda_1 R_1(x) + \lambda_2 R_2(x)$$

Possible choices are:

$$R_1(x) = \|Lx\|_1 \quad \text{and} \; R_2(x) = \|Lx\|_2^2$$

This is called l1-l2 regularization and combines the effects of both the priors

The regularization parameter

The value of the regularization parameter is the most critical setting in the model-based approach.

- A Too small value produeces a noisy reconstructed image, whereas a too large value produces an Image with extreme characteristics imposed by the regularization term.
- There are some *rules* for a good choice of the regularization parameter. However they usually requires multiple executions of the minimization problems to finally choose *the best one* for some predefined criteria. However this is too time consuming in imaging applications.
- The criteria can be split in:
    - ❖ Criteria using an estimate of the noise e
    - ❖ Criteria that do not use any infomration on the noise e

The regularization parameter

The general idea is that the solution corresponding to the regularization parameter λ should satisfy:

$$x_\lambda \to x^{exact} \qquad \text{as} \qquad \|e\|_2 \to 0$$

In particular is important to develop strategies whose convergence is as fast as possible.

The Discrepancy Principle (or Morozov principle)

The Discrepancy Principle is an *a posteriori criterion* which applies to an inverse problem modelled by:

$$Ax = y + e = y^\delta$$

Where e is white noise, i.e. random noise with normal distribution $N(0, \delta^2)$.

We suppose that the problem is ill-posed in the sense that the solution does not depend continuously on the data.

Discrepancy principle(DP).

Let $\tau \geq 1$ be a given number. Choose the regularization parameter. λ So that:

$$\|Ax_\lambda - y^\delta\|_2 = \tau\delta$$

N.B. The DP requires to know (or estimate) δ .

The Discrepancy Principle (or Morozov principle)

In the DP we choose the regularization parameter such that the residual norm is equal to

An a priori upper bound δ_ϵ for. $\|e\|_2$, i.e.:

$$\|Ax_\lambda - y\|_2 = \delta_\epsilon \quad \text{where} \quad \|e\|_2 \leq \delta_\epsilon$$

The Discrepancy principle is sensitive to variations in the estimate of the error norm.

Iterative regularization

Iterative regularization is based on the idea that when original least-squares problem:

$$min\|Ax - y^\delta\|_2^2$$

Is solved by means of an iterative method, we have the so called *semi-convergence* effect.

This means that in a iterations-error plot, the error curve has a convex shape, reaching its minumum and then Increasing again.

The regularization is obtained by stopping the iterations before the error increases, i.e. before the noise enters The reconstructed image.

Iterative regularization

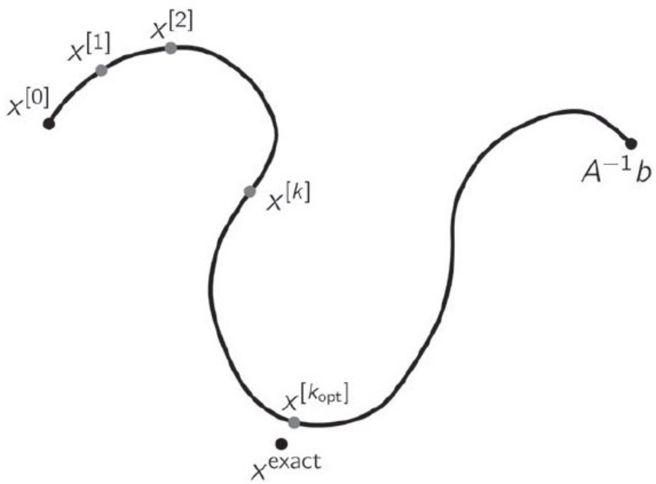


Figure 6.1. *The basic concept of semiconvergence. During the first iterations, the iterates $x^{[k]}$ tend to be better and better approximations to the exact solution x^{exact} (the optimal iterate is obtained at iteration k_{opt}), but at some stage they start to diverge again and instead converge toward the "naive" solution $A^{-1}b$.*