IMAGE DECONVOLUTION

M.Bertero
DISI – Università di Genova

''Inverse Problems: Computational Methods and Emerging Applications''
IPAM – October 16, 2003

Outline

1 - Introduction

In most applications of image deconvolution to the physical sciences, in particular microscopy and astronomy, accurate information about statistical properties of the noise is available. As a consequence, frequently used approaches to deconvolution are based on the Maximum Likelihood (ML) method. A particular case is the Least-Squares (LS) method which is the traditional starting point of the classical regularization theory.

It is known that the LS-formulation of deconvolution leads to ill-posed problems (ill-conditioned in their discrete approximation); even if no rigorous proof exists, it is a common opinion that all the existing ML-formulations of deconvolution are also leading to ill-posed problems; hence the need of regularization.

3

In most cases, iterative methods, converging to the ML-solutions, are used in such a way that regularization can be obtained by early stopping of the iterations. Indeed, these methods have the so-called *semi-convergence* property: the iterates first approach the "correct" solution and then go away.

However, due to the statistical setting of ML-methods, the most general appoach to regularization is provided by the so-called Bayesian methods.

In this tutorial, basic facts about image deconvolution are provided and the foundations of the ML formulations are given, as well as the iterative methods most frequently used in the applications. Moreover the Bayes approach and its relationship with the classical regularization theory are discussed.

4

2 - Mathematical modeling of image formation

An image is represented by a function of two or three variables $\underline{x}$:

$$\underline{x} \in \Re^n \ (n = 2, 3) \quad = \quad \text{coordinates in the image domain;}$$

space variables in the case of a microscope or a camera, angular variables in the case of a telescope (the case n = 3 is frequent in microscopy).

Notations :

$f(\underline{x}) =$ intensity distribution of the object to be imaged ;

$g_0(\underline{x}) =$ ideal image of $f(\underline{x})$ ($g_0(\underline{x})$ and $f(\underline{x})$ are defined in the same domain);

$g(\underline{x}) =$ detected image .

5

2.1 – Linear and space-invariant models

In most imaging systems the ideal image is approximately a linear function of the object; therefore the imaging system defines a linear operator A such that:

$$g_0 = A\, f .$$

Space-invariant systems are such that, if an object is translated by a, then the corresponding ideal image is also translated by a, i. e. the operator A commutes with the translation operators:

$$A(T_{\underline{a}}\, f) = T_{\underline{a}}\, g_0 = (T_{\underline{a}}\, A) f \ ,$$

$$(T_{\underline{a}}\, h)(\underline{x}) = h(\underline{x} - \underline{a}) .$$

6

A linear and continuous operator commuting with the translations is a convolution operator, i. e.there exists a function h(x) such that:

$$(Af)(\underline{x}) = \int h(\underline{x} - \underline{x}')\, f(\underline{x}')\, d\underline{x}' \ .$$

The function h(x) is called the *Point Spread Function* (PSF) of the imaging system: it is the image of a point source (the image of a point is not a point, as in geometric optics).

The ideal image is just a function in the range of the operator A which, in general, is assumed to be defined in a space of square integrable functions:

$$g_0(\underline{x}) = \int h(\underline{x} - \underline{x}')\, f(\underline{x}')\, d\underline{x}' \ .$$

We denote by H the Fourier transform of a function h:

$$H(\underline{w}) = \int h(\underline{x})\, e^{-i\,\underline{w}.\underline{x}}\, d\underline{x} \ .$$

From the convolution theorem we have the relation:

$$G_0(\underline{w}) = H(\underline{w})\, F(\underline{w}) \ .$$

The function H(ω) is called the *Transfer Function* (TF) of the imaging system.

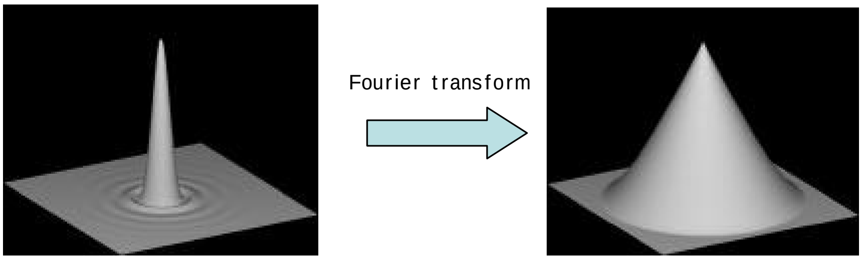
$$\underline{w} \in \Re^n \ (n = 2, 3) \ = \ \text{coordinates in frequency domain};$$

in the case of natural or microscopic images they are also called *space frequencies*, while in the case of astronomical images they are called *angular frequencies* and the two coordinates are usually denoted as u, v .

Example of diffraction-limited PSF in the case of a telescope consisting of a perfect circular mirror (diameter D) and monochromatic radiation (wavelength λ): the PSF is the so-called Airy function given by:

$$h(x,h) = \frac{\Omega^2}{p}\left(\frac{J_1(\Omega\,|J\,|)}{\Omega\,|J\,|}\right)^2 \quad , \quad |J\,| = \sqrt{x^2 + h^2} \quad , \quad \Omega = p\,\frac{\mathrm{D}}{l}$$

Similar structure in the case of an ideal microscope.



In most models the PSF satisfies the conditions:

i) $\quad h(\underline{x}) \geq 0$;

ii) $\quad \int h(\underline{x})\, d\underline{x} < \infty$.

The first property implies that the ideal image is non-negative if the object is non-negative while the second one implies that the imaging system is a low-pass filter in frequency domain. Indeed, from the Riemann-Lebesgue theorem it follows that the TF satisfies the conditions:

i) $H(\underline{w})$ is a bounded and continuous function ;

ii) $|H(\underline{w})| \rightarrow 0 \;,\; |\underline{w}| \rightarrow \infty$.

Very often the TF is zero outside a bounded domain (for instance a disc in 2D); in such a case the ideal images are band-limited functions, which can be represented by means of Whittaker-Shannon sampling expansions. 10

2.2 –Discrete models

Images are usually sampled on a uniform grid. In the case of bandlimited functions, the grid spacing is typically smaller, by a factor of 2 or 3, than the Nyquist-Shannon sampling distance, in order to avoid the need of interpolation of the measured data.

For simplicity we consider only 2D images. Then a sampling point is denoted by a pair of indices j,k taking values from 0 to N-1. More precisely the samples are the integrals over a small domain (pixel) and are are also called the pixel- values of the image (object). Then the indices j,k characterize the pixels of the domain. Since the object and the ideal image are represented on the same grid, we will denote their sampling values as $f[j,k]$ and $g_0[j,k]$, respectively.

11

If the PSF is negligible outside a domain with diameter D, then, from the support properties of the convolution product, it follows that the ideal image contains a complete representation of the object f if the object is zero in a strip of width D all around the boundary of the image domain. In such a case it is possible to approximate the continuous convolution by means of a discrete (cyclic) convolution:

$$g_0[j,k] = \sum_{j',k'=0}^{N-1} h[j-j',k-k']f[j',k'] \ .$$

This approximation is useful in the implementation of deconvolution algorithms which can be based on FFT. With an abuse of notation, in the following we will denote by A the block-circulant matrix with circulant blocks, approximating the convolution operator A, so that we write again:

$$g_0 = Af \ .$$

12

We will use again capital letters for denoting the Discrete Fourier Transform (DFT) of an array, i. e. if the array is h[j,k], its DFT is denoted by H[l,m] and is given by:

$$H[l,m] = \sum_{j,k=0}^{N-1} h[j,k] e^{-i\frac{2p}{N}(l\,j + m\,k)} \ .$$

Then, from the convolution theorem for the discrete convolution product, we have again:

$$G_0[l,m] = H[l,m]F[l,m] \ .$$

The arrays h[j,k] and H[l,m] can be called, respectively, the discrete PSF and the discrete TF.

13

2.3 – Noise

In the following we consider only discrete models. We denote by g the array formed by the detected values, g[j,k], of an image. The detected values differ from the values of the ideal image as a consequence of several perturbing effects. The most important ones are the background and the noise. The background is due to a diffuse emission of the surrounding medium, which can be assumed approximately constant over the image domain. The noise is due to the detection system and is random. In general one can write:

$$g[j,k] = g_0[j,k] + b + n[j,k] \ ,$$

where b denotes the constant background and n[j,k] the noise contribution (in general, not additive).

14

Due to the randomness of noise, it is quite natural to consider the detected values g[j,k] as realizations of random variables (r.v.) which will be denoted by η[j,k]. Therefore we denote by η the array formed by these random variables; the realization of η is just g.

For simplicity we assume that, for any given f, all r.v. s η[j,k] are absolutely continuous so that they are described by a joint probability density denoted by $P_\eta(g|f)$.

The following properties are assumed in most models:

i) the probability density $P_\eta(g\,|\,f)$ can be derived from the noise properties ;

ii) the following relationship holds true :

$$E[\eta\,|\,f] = Af + b \quad ,$$

where $E[\eta\,|\,f]$ denotes the expected value of η given f.

15

Model 1 – Additive noise

In some cases the noise due to the detection device (in particular the so-called read-out-noise, RON, due to the electronics of the CCD camera) is the realization of r.v. s which are independent of the object f. If we denote by ν the array formed by the noise r.v. s ν[j,k], described by the joint probability density $P_\nu(n)$, from the relatioship:

$$\eta = Af + b + \nu \quad ,$$

we obtain that the probability density of η , given f, is related to the probability density of ν by the equation:

$$P_\eta(g|f) = P_\nu(g - Af - b) \; .$$

Gaussian noise is an important particular case.

16

If the r. v. s $\nu[j,k]$ are Gaussian with zero expected values and covariance matrix C, then :

$$P_\nu(\mathrm{n}) = \left(\frac{1}{2\pi\,|C|}\right)^N \exp\left\{-\frac{1}{2}\left\langle \mathrm{n}, C^{-1}\mathrm{n}\right\rangle_2\right\} \ ,$$

where |C| denotes the determinant of C and $\langle \cdot\,,\cdot\rangle_2$ is the usual Euclidean scalar product (we assume images N xN).

In the case of the so-called white noise the r.v. s $\nu[j,k]$ are independent, and have the same standard deviation σ :

$$P_\nu(\mathrm{n}) = \left(\frac{1}{2\pi\sigma^2}\right)^N \exp\left\{-\frac{1}{2\sigma^2}\|\mathrm{n}\|_2^2\right\} ,$$

so that :

$$P_\eta(\mathrm{g}|\mathrm{f}) = \left(\frac{1}{2\pi\sigma^2}\right)^N \exp\left\{-\frac{1}{2\sigma^2}\|\mathrm{g} - A\mathrm{f} - \mathrm{b}\|_2^2\right\} .$$

17

Model 2 – Poisson noise

In some applications, such as microscopy and astronomy, the measurements are based on countings of events (photons); in these cases the signals can be modeled as Poisson processes. The r. v. s $\eta[j,k]$ can, in general, be assumed as independent (integer valued) r. v. s, with expected values given by assumption ii). Since the detected values g[j,k] are integer numbers, we can write:

$$P_\eta(\mathrm{g}|\mathrm{f}) = \prod_{j,k=0}^{N-1} \frac{\left[(A\,\mathrm{f})[j,k] + \mathrm{b}\right]^{\mathrm{g}[j,k]}}{\mathrm{g}[j,k]\,!}\, e^{-\left[(A\,\mathrm{f})[j,k]+\mathrm{b}\right]} \ .$$

Remember that, in the case of a Poisson r. v., the expected value coincides with the variance.

18

Model 3 – Poisson and Gaussian noise

More generally the measurement process is a Poisson counting process contaminated by additive (independent) Gaussian noise. In such a case the statistical model for the detected image values is as follows [3]:

$$P_\eta(g|f) =$$

$$= \prod_{j,k=0}^{N-1} \sum_{m=0}^{+\infty} \frac{\left[(Af)[j,k]+b\right]^m}{m!} e^{-[(Af)[j,k]+b]} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2\sigma^2} \|g[j,k]-m\|_2^2 \right\},$$

where we have considered the particular case of white noise with zero expected value.

If needed, this expression can be simplified by using approximations (for instance Stirling formula).

19

3 – Maximum likelihood methods

In classical statistics Maximum Likelihood (ML) is the most commonly used method for parameter estimation. Its application to image restoration is based on the knowledge of the random properties of noise discussed in Section 1.3, so that the probability density $P_\eta(g|f)$ is known. Then the ML estimator answers to the following question:

Which object f is most likely to produce the detected image g ?

Definition 1 – For a given detected image g the likelihood function is the function of the object f defined by :

$$L_g(f) = P_\eta(g|f) \; .$$

20

Definition 2 – A ML-estimate of the object f is any object which maximizes the likelihood function:

$$f_{ML} = \arg\max_f L_g(f) \ .$$

Remark – In the case of image deconvolution, this problem is, in general, ill-posed.

Since in all practical applications the probability density of η is the product of a large number of factors, it is useful to introduce the following log-likelihood function:

$$l_g(f) = \ln L_g(f) \ ,$$

so that the ML-estimate is also given by:

$$f_{ML} = \arg\max_f l_g(f) \ .$$

21

Relationships with the usual approach to Inverse Problems can be clarified by considering functionals of the following type:

$$J_g(f) \propto -\ln l_g(f) \ ,$$

obtained by dropping or adding terms depending only on g, so that the ML estimate is also given by:

$$f_{ML} = \arg\min_f J_g(f) \ .$$

In all ML approaches it is possible to introduce additional constraints on the object such as non-negativity, bounded support of the solution etc. According to the previous formulation all these problems are equivalent to problems of constrained minimization.

22

3.1 – The Gaussian case

In the case of additive Gaussian noise the likelihood function is given in Section 1.3, Model 1. It follows that:

$$J_g(f) = \left\langle g - Af - b, C^{-1}\left(g - Af - b\right)\right\rangle_2 \ ,$$

and the ML-method is equivalent to the least-squares method. In the particular case of white noise:

$$J_g(f) = \| g - Af - b \|_2^2 \ .$$

Any minimum point of this functional is called a *least-squares* (LS) *solution*:

$$f_{LS} = \arg\min_f J_g(f) \ ,$$

and is a solution of the Euler equation:

$$A^T C^{-1} A f_{LS} = A^T C^{-1} (g - b) \ .$$

23

If the matrix A is a block circulant matrix, as discussed in Section 1.2, then, by introducing the "subtracted" image:

$$g_s = g - b \ ,$$

and by taking the DFT of both sides of the Euler equation in the case C = I (white noise), we obtain:

$$| H[l, m] |^2 F_{LS}[l, m] = H^*[l, m] G_s[l, m] \ ,$$

where the * denotes complex conjugation. Therefore, in the points where H[k,l] = 0, the value of the FT of the LS solution is not determined, while in the other points it is given by:

$$F_{LS}[l, m] = \frac{G_s[l, m]}{H[l, m]} \ .$$

24

The least-squares problem is ill-posed because the solution may not be unique and is strongly corrupted by noise propagation (ill-conditioning of the imaging matrix). The generalized solution, or minimal norm LS solution, is the LS solution with minimal Euclidean norm. It is obtained by setting:

$$F_{LS}[l, m] = 0 \quad \text{if} \quad H[l, m] = 0 \ ,$$

and therefore it is given by:

$$f_{LS}[j,k] = \frac{1}{N^2} \sum_{H[l,m]\neq 0} \frac{G_s[l,m]}{H[l,m]} e^{-i\frac{2p}{N}(j\,l + k\,m)} \ .$$

The numerical instability of this solution is due to the small values of H[l,m] which amplify the noise terms of the corresponding Fourier components of the subtracted image.

25

The numerical instability of the generalized solution is related to the uncertainty on the solutions which are compatible with the data within a given noise level. The set of these solutions can be reduced by means of constraints. The most popular is non-negativity. It leads to the definition of non-negative least-squares solutions:

$$f_{LS}^{+} = \arg\ \min_{f \geq 0} \ \| A f - g_s \|_2 \ .$$

Non-negativity, in general, does not remove the numerical instability of the solution nor the non-uniqueness (it is easy to provide examples). However, iterative algorithms have been proposed, converging to the non-negative LS solutions and having the semi-convergence property: first the iterates approach the correct solution and then go away.

26

One of these methods is that called ISRA (Iterative Space Reconstruction Algorithm), whose iterations are defined by:

i) give $f^{(0)} \geq 0$;

ii) given $f^{(k)}$, compute :

$$f^{(k+1)} = f^{(k)} \frac{A^T g_s}{A^T A f^{(k)}} ,$$

where the quotient is the image obtained by dividing the two images pixel by pixel. The computation of each iteration is fast since it basically requires two FFT s for the computation of the denominator. The convergence of the algorithm to non-negative LS–solutions for any given initial guess has been proved by De Pierro [4]. The semi-convergence of the algorithm is an experimental result derived from numerical simulations.

27

The second method is that called Projected Landweber method, with iterations defined by:

i) give $f^{(0)} \geq 0$;

ii) given $f^{(k)}$, compute :

$$f^{(k+1)} = P_+ \left\{ f^{(k)} + t \left(A^T g_s - A^T A f^{(k)} \right) \right\} ,$$

$$0 < t < \frac{2}{H_{max}^2} ,$$

where P+ is the projection on the cone of non-negative images and Hmax is the maximum value of |H[l,m]|. Again the implementation of each iteration requires the computation of two FFT s. The convergence to non- negative LS solutions has been proved by Eicke [5]. The semi-convergence is proved in the non-projected case and is an experimental result in the projected one. Moreover the metod applies to any convex constraint.

28

2.2 – The Poisson case

In the case of Poisson noise, Model 2 of Section 1.3, the ML method leads to the minimization of the following functional, which is obtained from the log-likelihood function by modifying the g-dependent term:

$$J_g(f) = \sum_{j,k=0}^{N-1} \left\{ g[j,k] \ln \frac{g[j,k]}{(A\,f)[j,k]+b} + \left[(A\,f)[j,k]+b - g[j,k] \right] \right\}.$$

This functional can be interpreted as a directed distance between g and Af + b and is just the Csiszar I-divergence measure of the discrepancy between g and Af + b. It is defined for all f such that Af + b > 0. If the PSF is non-negative then the natural domain of the functional is the closed cone of the non-negative f.

29

The functional $J_g(f)$ has the following properties :

i) $J_g(f) \geq 0$;

ii) $J_g(f)$ is convex.

Property i) follows from the elementary inequality:

$$a \ln a - a \ln x + x - a \geq 0 \;,$$

for fixed a > 0 and any x > 0.

As concerns property ii), we need to compute the gradient and the Hessian of the functional; in order to simplify the equations, we assume that the PSF satisfies the condition:

$$\sum_{j,k=0}^{N-1} K[j,k] = 1 \;,$$

which can always be obtained by a suitable normalization of the PSF.

30

The gradient of the functional is given by:

$$\frac{\partial J_g(f)}{\partial f\,[j,k]} = - \sum_{j',k'=0}^{N-1} h[j'-j,k'-k]\frac{g[j',k']}{(A\,f)[j',k']+b} + \sum_{j',k'=0}^{N-1} h[j'-j,k'-k] =$$

$$= -\left(A^T \frac{g}{A\,f+b}\right)[j,k] + 1\ ,$$

where :

$$(\,A^T\ f\,)\,[j,k]\ =\ \sum_{j',k'=0}^{N-1} h[j'-j,k'-k]\,f\,[j',k']\ ,$$

and the quotient is defined again by pixel by pixel division. Similar definition for the product of two images. As concerns the Hessian we have:

$$\frac{\partial^2\,J_g\,(f)}{\partial f\,[j,k]\,\partial f\,[j',k']}\ =\ \left[H_g\,(f)\right][j,k\,;\,j',k']=$$

31

$$=\ \sum_{j'',k''=0}^{N-1} h[j''-j,k''-k]\frac{g\,[j'',k'']}{\left[(A\,f)[j'',k'']+b\right]^2} h[j''-j',k''-k']\ ,$$

and it is positive semi-definite since, for any array u :

$$\sum_{j,k=0}^{N-1}\ \sum_{j',k'=0}^{N-1} \left[H_g\,(f)\right][j,k\,;\,j',k']\ \ u[j,\,k]\ \ u[j',k']\ =$$

$$=\ \sum_{j'',k''=0}^{N-1} g\,[j'',k'']\left\{\frac{(\,A\,u\,)\,[j'',k'']}{(\,A\,f\,)\,[j'',k'']\ +\ b}\right\}^2\ \ \geq\ 0\ .$$

It follows that the functional is convex. The Hessian is positive definite if and only the equation Au = 0 implies u = 0 ; in such a case the functional is strictly convex and therefore a minimum point, if it exists, is unique. This condition, however, is not satisfied in the case of band-limited imaging systems.

32

The previous properties imply that all minima are global.

Sufficient conditions are the Kuhn-Tucker conditions: a point f_{min} is a minimum point if:

$$f \frac{\partial J_g(f)}{\partial f}\bigg|_{f=f_{min}} = 0 \ ,$$

$$\frac{\partial J_g(f)}{\partial f[j,k]}\bigg|_{f=f_{min}} \leq 0 \ , \quad \text{if} \ \ f_{min}[j,k] = 0 \ .$$

The first condition implies:

$$f = f \, A^T \left(\frac{g}{A f + b} \right) ,$$

which can be formally treated as a fixed point equation.

33

By applying the method of successive approximations to the previous fixed point equation, we obtain the iterative algorithm:

i) give $f^{(0)} \geq 0$;

ii) given $f^{(k)}$, compute :

$$f^{(k+1)} = f^{(k)} \, A^T \frac{g}{A \, f^{(k)} + b} \ .$$

The computation of one iteration of this method is more expensive than the computation of one iteration of ISRA or of the projected Landweber method; indeed it requires the computation of four FFT s: two for computing the denominator and two for computing the result of the application of the transposed matrix. In all cases the convergence (or semi-convergence) is slow: it requires the computation of many iterations.

34

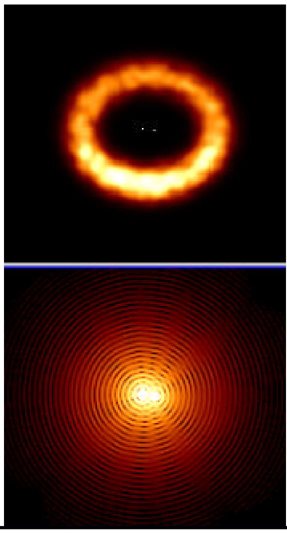
It is obvious that all the iterates are non-negative if the image g and the PSF h are also non-negative. Moreover, in the case b = 0, they satisfy the condition:

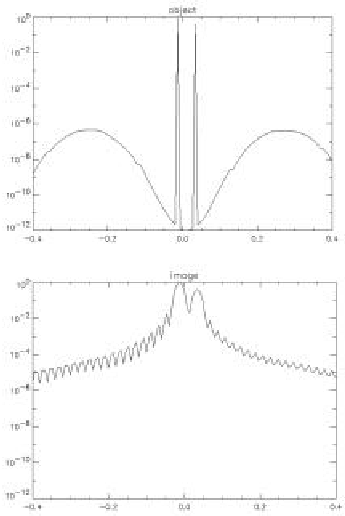
$$\sum_{j,k=0}^{N-1} f^{(k)}[j,\ k] \ = \ \sum_{j,k=0}^{N-1} g[j,\ k]\ .$$

This method has been introduced (independently) by Richardson [6] and Lucy [7] and, for this reason, is known as Richardson-Lucy (RL) [6,7] method in Astronomy (it was used for the deconvolution of the aberrated images of the Hubble Space Telescope). It was rediscovered ten years later by Shepp and Vardi [8] as a method for Emission Tomography. They proved that it is a particular case of the Expectation Maximization (EM) method for the solution of ML problems and they also proved (in a weak sense) that it converges to maximum points of the likelihood function. More complete proofs are given in [9-10]. A generalization to a continuous model is given in [11].

35

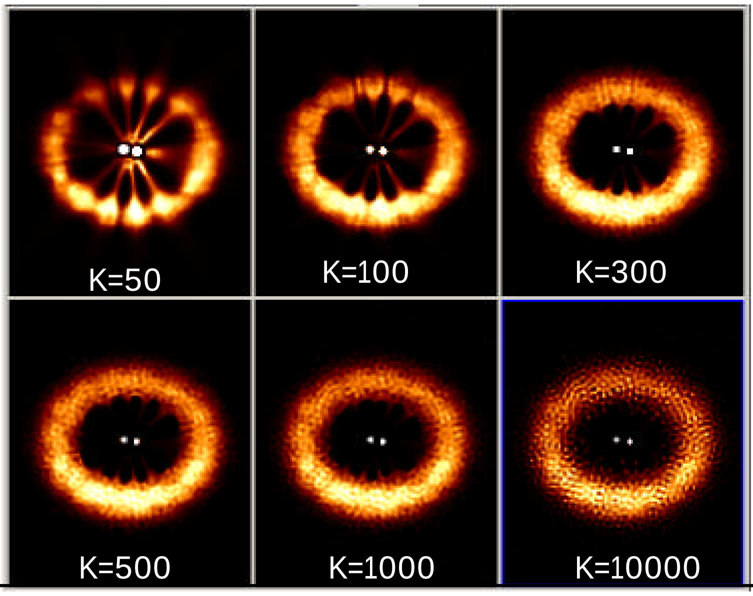
Example of an astronomical object (a model of young binary star) and of its image.





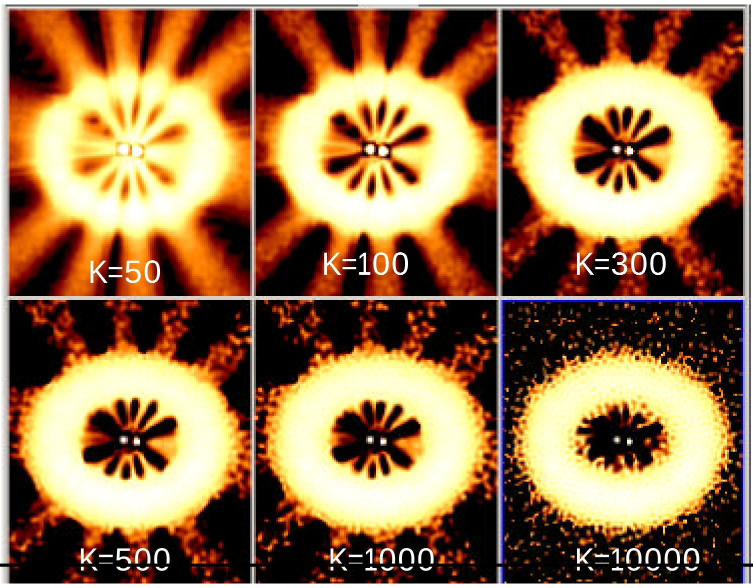
36

RL-iterates in the case of the previous image (linear scale)



37

RL-iterates in the case of the previous image (log-scale)



38

4 – Bayesian methods and regularization

The ill-posedness (ill-conditioning) of ML-problems is generated by a lack of information on the object f, in particular by the lack of information about the frequencies corresponding to small values of the TF H[l,m]. A remedy can be the use of additional (prior) information on f, which can be expressed in the form of constraints (for instance, non-negativity) on f or of statistical properties of f.

In a probabilistic approach it is assumed that the object f and the image g are realizations of arrays of r.v. s, denoted respectively by ξ and η, and that the problem is solved if we know their joint probability density $P_{\xi\eta}(f,g)$.

39

As discussed in Chapter 2, the conditional probability density of η given f can be deduced from known statistical properties of the noise. However the marginal probability density of ξ, $P_{\xi}(f)$, in general is not known. One can guess this probability density, using his knowledge or ignorance about f. The model used is usually called a "prior".

If the marginal distribution of f is given, then one can obtain the joint probability density from Bayes formula :

$$P_{\xi\eta}(f,g) = P_{\eta}(g|f)\,P_{\xi}(f) \ .$$

Using the analogous Bayes formula for the conditional probability of ξ, given g, one obtains:

$$P_{\xi}(f|g) = \frac{P_{\eta}(g|f)\,P_{\xi}(f)}{P_{\eta}(g)} \ ,$$

which is the basic tool in the so-called Bayesian methods.

40

A Maximum A Posteriori (MAP) of the object f, given the image g, is any solution of the problem:

$$f_{MAP} = \arg\max_f \; P_x(f|g) \; .$$

Introducing also in this case the log-function and neglecting the term independent of f, one finds:

$$f_{MAP} = \arg\max_f \; \left\{ l_g(f) + \ln P_x(f) \right\} .$$

In terms of the functionals $J_g(f)$ introduced in Chapter 2, the problem becomes:

$$f_{MAP} = \arg\;\min_f \left\{ J_g(f) - \ln P_x(f) \right\} \; ,$$

and therefore the term $-\ln P_x(f)$ looks as a regularization of the functional $J_g(f)$.

41

The most frequently used priors are of the Gibbs type, namely of the form:

$$P_x(f) = C \; \exp \; \left\{ -m \, \Omega(f) \right\} ,$$

where μ is a parameter, which can play the role of the regularization parameter, and Ω (f) is a functional expressing prior information about the object to be estimated.

In such a case, the MAP problem takes the following form:

$$f_{MAP} = \arg\min_f \; \left\{ J_g(f) + m \, \Omega(f) \right\} ,$$

and therefore it has the same structure of typical problems in regularization theory.

42

Examples of Gibbs priors:

i) $\Omega(f) = \| f \|_2^2$ (white noise prior);

ii) $\Omega(f) = \| L f \|_2^2$, where L is a discrete approximation of a differential operator, for instance the Laplacian (smoothness prior);

iii) $\Omega(f) = \| f \|_1$ (impulse noise);

iv) $\Omega(f) = TV(f)$ (total variation prior).

In the case of Gaussian white noise we obtain functionals of the classical regularization theory:

$$R_{g,m}(f) = \| A f - g_s \|_2^2 + m\Omega(f).$$

Therefore all functional analytic methods developed for this theory [2] apply to the investigation of MAP solutions in this case.

43

In the case of Poisson noise, the MAP estimates are obtained by minimizing the functional:

$$J_g(f) + m\ \Omega(f) =$$

$$= \sum_{j,k=0}^{N-1} \left\{ g[j,k] \ln \frac{g[j,k]}{(A f)[j,k] + b} + \left[(A f)[j,k] + b - g[j,k] \right] \right\} + m\,\Omega(f).$$

A complete theory has not yet been developed, even if several partial results are contained in the scientific literature. It is obvious that the functional is defined on the cone of non-negative f.

Several iterative methods have been proposed for the minimization of these functionals for various kinds of priors. A unified approach to most of these methods can be found in [12] .

44

5 - References

[1] M. Bertero and P. Boccacci, 1998, Introduction to Inverse Problems in Imaging, IOP, Bristol

[2] H. W. Engl, M. Hanke, and A. Neubauer, 1996, Regularization of Inverse Problems, Kluver, Dordrecht

[3] D. L. Snyder, A. M. Hammoud, and R. L. White, 1993, Image recovery from data acquired with a charged-coupled -device camera, 1993, J. Opt. Soc. Am., A-10, 1014-1023

[4] A. R. De Pierro, 1987, On the convergence of the iterative image space reconstruction algorithm for volume ECT, IEEE Trans. Med. Imaging, 6, 124-125

[5] B. Eicke, 1992, Iteration methods for convexly constrained ill-posed problems in Hilbert space, Num. Funct. Ana. Optim., 13, 413-429

45

[6] W. H. Richardson, 1972, Bayesian-based iterative method of image restoration, J. Opt. Soc. Am., 62, 55-59

[7] L. Lucy, 1974, An iterative technique for the rectification of observed distribution, Astron. J., 79, 745-754

[8] L. A. Shepp and Y. Vardi, 1982, Maximum likelihood reconstruction for emission tomography, IEEE Trans. Med. Imaging, 1, 113-122

[9] K. Lange, and R. Carson, 1984, EM reconstruction algorithm for emission and transmission tomography, J. Comp. Assisted Tomography, 8, 306-316

[10] Y. Vardi, L. A. Shepp, and L. Kaufman, 1985, A

46

statistical model for Positron Emission Tomography, J. Am. Stat. Asociation, 80, 8-20

[11] H. N. Multhei, and B. Schorr, 1987, On an iterative method for a class of integral equations of the first kind, Math. Meth. in the Appl. Sci., 9, 137-168

[12] H. Lanteri, M. Roche, and C. Aime, 2002, Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms, Inverse Problems, 18, 1397-1419