On the stopping criterion for numerical methods for linear systems with additive Gaussian noise

Rapport technique EPM-RT-2009-10
de l'École Polytechnique de Montréal

Benoit Hamelin,[*†] Yves Goussard[†] and Jean-Pierre Dussault[‡]

December 4th, 2009

[*]Corresponding author: benoit-2.hamelin@polymtl.ca

[†]Benoit Hamelin and Yves Goussard are with the Institut de génie biomédical, École Polytechnique de Montréal, Montréal, Québec, Canada.

[‡]Jean-Pierre Dussault is with the Département d'informatique, Université de Sherbrooke, Sherbrooke, Québec, Canada.

Abstract

We consider the inversion of a linear operator with centered Gaussian white noise by MAP estimation with a Gaussian prior distribution on the solution. The actual estimator is computed approximately by a numerical method. We propose a relation between the stationarity measure of this approximate solution to the mean square error of the exact solution. This relation enables the formulation of a stopping test for the numerical method, met only by iterates that satisfy chosen statistical properties. We extend this development to Gibbs priors using a quadratic extrapolation of the log-likelihood maximized by the MAP estimator.

1 Introduction

We consider the inversion of a system of equations where a Gaussian process is superposed to the image of a linear operator. Problems of this kind appear in many digital imaging and signal processing applications, among many other engineering domains. The inversion is naturally computed by means of a regularized estimator, which entails the resolution of a least squares minimization problem.

The exact formulation of this problem depends on the prior distribution that is assumed over the domain of the operator to invert. On the one hand, when this distribution is Gaussian, the least squares objective function is compounded to a quadratic penalization term, yielding a quadratic program. The solution of this program corresponds to the root of the normal equation, which is usually found by a direct method, such as QR decomposition, or by a sequential recurrent process. The latter approach is practical even in cases where the linear operator is not represented in an explicit matrix form. However, for large-scale problems, the full sequential process can prove prohibitively costly in terms of runtime. One must then turn to methods for symmetric definite positive linear systems, such as the conjugate gradient or Lanczos iterations (Kelley, 1995), which may converge to a good estimator at the cost of fewer operator evaluations. On the other hand, the use of a non-Gaussian prior distribution[1] (Idier, 2001) compounds a nonquadratic, though possibly smooth or convex, penalization term to the least-squares objective. Numerical methods for nonlinear optimization are then the only feasible approach for the task at hand.

In any such application where numerical methods are involved, the question of the stopping test is seldom addressed in full. In the Gaussian prior case, some norm of the residual of the normal equation is typically used to determine convergence. In the alternate case, nonlinear solvers all seem to have their own convergence criterion, which relies on measures of objective descent or stationarity. In both cases, the choice of the stopping tolerance for the convergence test is of significant incidence to the statistical properties of the estimator. Indeed, it is known that early termination of the numerical method entails implicit regularization of the estimator that is equivalent to adding a supplemental quadratic penalization term to the objective being minimized (Engl et al., 1996). This is unacceptable when using a non-Gaussian prior distribution, as it ruins the specific properties of the estimator brought by the nonquadratic penalization. For a Gaussian prior, the supplemental penalization inflates the bias of the estimator, as it is positively correlated to the weight of the penalization terms of the least-squares objective. The opposite error with respect to the choice of the stopping tolerance raises some practical issues, as it augments the runtime of the estimation, possibly past acceptable levels.

We report herein on a relation between the mean square error (MSE) of an estimator obtained from a truncated numerical method and a stationarity measure of this estimator with respect to the least-squares problem. We deduct from this relation a statistical interpretation to the stopping tolerance, which may then be chosen with respect to a tolerated error level.

The rest of this paper is as follows. Section 2 recalls the properties of the maximum *a*

[1] The nonquadratic penalization brought by the use of a non-Gaussian prior distribution on the solution favors certain desirable properties of the solution, such as the better preservation of high-frequency structures with respect to quadratic penalization.

posteriori (MAP) estimator in the case of a Gaussian prior distribution. Section 3 discusses a variant of this estimator that corresponds to an approximate solution to the least squares problem and derives a stopping criterion for numerical methods in the Gaussian prior case. Section 4 extends these results to the case of a non-Gaussian prior distribution. Section 5 offers some concluding remarks.

2 MAP estimation with a Gaussian prior distribution

Let us first recall the inverse problem at hand. We have[2]

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x} + \boldsymbol{b},$$

with $\boldsymbol{y} \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, $\mathbf{A}$ a matrix of dimensions $m \times n$ and

$$\boldsymbol{b} \rightsquigarrow \mathrm{N}\left(0, \sigma^2 I_m\right).$$

The distribution of $\boldsymbol{y}$ conditional on $\boldsymbol{x}$ is then Gaussian of expectation $A\boldsymbol{x}$ and covariance $\sigma^2 I_m$, so it is expressed as

$$f_{\boldsymbol{y}}(\boldsymbol{y} \mid \boldsymbol{x}) = \frac{1}{(2\pi\sigma^2)^{m/2}} \exp\left(-\frac{(\boldsymbol{y} - \mathbf{A}\boldsymbol{x})^{\mathrm{t}}(\boldsymbol{y} - \mathbf{A}\boldsymbol{x})}{2\sigma^2}\right).$$

We also set the prior distribution of $\boldsymbol{x}$ as Gaussian of null expectation and covariance $\sigma_x^2(\mathbf{R}^{\mathrm{t}}\mathbf{R})^{-1}$, so its density is

$$f_{\boldsymbol{x}}(x) = \frac{|\mathbf{R}^{\mathrm{t}}\mathbf{R}|}{(2\pi\sigma_x^2)^{n/2}} \exp\left(-\frac{\boldsymbol{x}^{\mathrm{t}}\mathbf{R}^{\mathrm{t}}\mathbf{R}\boldsymbol{x}}{2\sigma_x^2}\right).$$

This model allows for the estimation of $\boldsymbol{x}$ by maximizing the log-likelihood of the posterior density of $\boldsymbol{x}$ with respect to the given $\boldsymbol{y}$. This MAP estimator solves the least-squares optimization problem

$$\min \frac{1}{2}\|\boldsymbol{y} - \mathbf{A}\boldsymbol{x}\|^2 + \frac{\lambda}{2}\|\mathbf{R}\boldsymbol{x}\|^2, \tag{1}$$

with $\lambda = \sigma^2/\sigma_x^2$. It therefore satisfies

$$(\mathbf{A}^{\mathrm{t}}\mathbf{A} + \lambda\mathbf{R}^{\mathrm{t}}\mathbf{R})\hat{\boldsymbol{x}} = \mathbf{N}\hat{\boldsymbol{x}} = \mathbf{A}^{\mathrm{t}}\boldsymbol{y}, \tag{2}$$

with $\mathbf{N} = \mathbf{A}^{\mathrm{t}}\mathbf{A} + \lambda\mathbf{R}^{\mathrm{t}}\mathbf{R}$, so we can write

$$\hat{\boldsymbol{x}} = \mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}. \tag{3}$$

The minimal mean square error (conditional to $\boldsymbol{y}$) estimator is known to be

$$\hat{\boldsymbol{x}}_{\mathrm{MMSE}} = \mathrm{E}\left[\boldsymbol{x} \mid \boldsymbol{y}\right]$$

[2]In the following a random variable and its realizations will be noted samewise, unless where it would be confusing.

(Shanmugan and Breipohl, 1988). In this case, we may show that it corresponds to the MAP estimator (3) by expressing the probability density of $\boldsymbol{x}$ conditional on $\boldsymbol{y}$:

$$f_{\boldsymbol{x}}(\boldsymbol{x} \mid \boldsymbol{y}) = \frac{f_{\boldsymbol{y}}(\boldsymbol{y} \mid \boldsymbol{x}) f_{\boldsymbol{x}}(\boldsymbol{x})}{\int_{-\infty}^{\infty} f_{\boldsymbol{y}}(\boldsymbol{y} \mid \boldsymbol{x}) f_{\boldsymbol{x}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}}. \tag{4}$$

Let us develop equation (4) starting with its numerator.

$$\begin{aligned}
f_{\boldsymbol{y}}(\boldsymbol{y} \mid \boldsymbol{x}) f_{\boldsymbol{x}}(\boldsymbol{x}) = {} & \frac{1}{(2\pi)^{n/2}|\sigma_x^2(\mathbf{R}^{\mathrm{t}}\mathbf{R})^{-1}|^{1/2}(2\pi)^{m/2}|\sigma^2 I|^{1/2}} \times \\
& \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{y} - \mathbf{A}\boldsymbol{x})^{\mathrm{t}}(\boldsymbol{y} - \mathbf{A}\boldsymbol{x}) - \frac{1}{2\sigma_x^2}\boldsymbol{x}^{\mathrm{t}}\mathbf{R}^{\mathrm{t}}\mathbf{R}\boldsymbol{x}\right) \\
= {} & \frac{1}{(2\pi)^{n/2}|\sigma_x^2(\mathbf{R}^{\mathrm{t}}\mathbf{R})^{-1}|^{1/2}(2\pi)^{m/2}|\sigma^2 I|^{1/2}} \exp\left(-\frac{\boldsymbol{y}^{\mathrm{t}}\boldsymbol{y}}{2\sigma^2}\right) \times \\
& \exp\left(-\frac{-\boldsymbol{y}^{\mathrm{t}}\mathbf{A}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{t}}\mathbf{A}^{\mathrm{t}}\boldsymbol{y} + \boldsymbol{x}^{\mathrm{t}}\mathbf{N}\boldsymbol{x}}{2\sigma^2}\right) \\
= {} & \frac{1}{(2\pi)^{n/2}|\sigma_x^2(\mathbf{R}^{\mathrm{t}}\mathbf{R})^{-1}|^{1/2}(2\pi)^{m/2}|\sigma^2 I|^{1/2}} \exp\left(-\frac{\boldsymbol{y}^{\mathrm{t}}\boldsymbol{y} + \boldsymbol{y}^{\mathrm{t}}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}}{2\sigma^2}\right) \times \\
& \exp\left(-\frac{\boldsymbol{x}^{\mathrm{t}}\mathbf{N}\boldsymbol{x} - \boldsymbol{y}^{\mathrm{t}}\mathbf{A}\mathbf{N}^{-1}\mathbf{N}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{t}}\mathbf{N}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y} + \boldsymbol{y}^{\mathrm{t}}\mathbf{A}\mathbf{N}^{-1}\mathbf{N}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}}{2\sigma^2}\right) \\
= {} & \frac{1}{(2\pi)^{n/2}|\sigma_x^2(\mathbf{R}^{\mathrm{t}}\mathbf{R})^{-1}|^{1/2}(2\pi)^{m/2}|\sigma^2 I|^{1/2}} \exp\left(-\frac{\boldsymbol{y}^{\mathrm{t}}\boldsymbol{y} + \boldsymbol{y}^{\mathrm{t}}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}}{2\sigma^2}\right) \times \\
& \exp\left(-\frac{(\boldsymbol{x} - \mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y})^{\mathrm{t}}\mathbf{N}(x - \mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y})}{2\sigma^2}\right).
\end{aligned} \tag{5}$$

Therefore, the denominator of equation (4) may be expressed as

$$\begin{aligned}
f_{\boldsymbol{y}}(\boldsymbol{y}) = {} & \int_{-\infty}^{\infty} f_{\boldsymbol{y}}(\boldsymbol{y} \mid \boldsymbol{x}) f_{\boldsymbol{x}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \\
= {} & \frac{1}{(2\pi)^{n/2}|\sigma_x^2(\mathbf{R}^{\mathrm{t}}\mathbf{R})^{-1}|^{1/2}(2\pi)^{m/2}|\sigma^2 I|^{1/2}} \exp\left(-\frac{\boldsymbol{y}^{\mathrm{t}}\boldsymbol{y} + \boldsymbol{y}^{\mathrm{t}}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}}{2\sigma^2}\right) \times \\
& \int_{-\infty}^{\infty} \exp\left(-\frac{(\boldsymbol{x} - \mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y})^{\mathrm{t}}\mathbf{N}(x - \mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y})}{2\sigma^2}\right) \mathrm{d}\boldsymbol{x}.
\end{aligned} \tag{6}$$

Let $u = \frac{1}{\sigma}\mathbf{N}^{1/2}(\boldsymbol{x} - \hat{\boldsymbol{x}})$, for which $\nabla\boldsymbol{x}(u) = \sigma N^{-1/2}$; then equation (6) resolves to

$$\begin{aligned}
f_{\boldsymbol{y}}(\boldsymbol{y}) = {} & \frac{|\sigma\mathbf{N}^{-1/2}|}{(2\pi)^{n/2}|\sigma_x^2(\mathbf{R}^{\mathrm{t}}\mathbf{R})^{-1}|^{1/2}(2\pi)^{m/2}|\sigma^2 I|^{1/2}} \exp\left(-\frac{\boldsymbol{y}^{\mathrm{t}}\boldsymbol{y} + \boldsymbol{y}^{\mathrm{t}}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}}{2\sigma^2}\right) \times \\
& \int_{-\infty}^{\infty} \mathrm{e}^{-u^{\mathrm{t}}u/2} \mathrm{d}u \\
= {} & \frac{|\sigma\mathbf{N}^{-1/2}|(2\pi)^{m/2}}{(2\pi)^{n/2}|\sigma_x^2(\mathbf{R}^{\mathrm{t}}\mathbf{R})^{-1}|^{1/2}(2\pi)^{m/2}|\sigma^2 I|^{1/2}} \exp\left(-\frac{\boldsymbol{y}^{\mathrm{t}}\boldsymbol{y} + \boldsymbol{y}^{\mathrm{t}}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}}{2\sigma^2}\right).
\end{aligned} \tag{7}$$

By putting (5) and (7) together, we get

$$f_{\boldsymbol{x}}(\boldsymbol{x} \mid \boldsymbol{y}) = \frac{1}{(2\pi)^{n/2}\,|\sigma^2\mathbf{N}^{-1}|} \exp\left(-\frac{(\boldsymbol{x}-\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y})^{\mathrm{t}}\mathbf{N}(\boldsymbol{x}-\mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y})}{2\sigma^2}\right). \tag{8}$$

Equation (8) effectively describes a Gaussian probability density, with

$$\begin{aligned}\mathrm{E}\left[\boldsymbol{x} \mid \boldsymbol{y}\right] &= \mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y} = \hat{\boldsymbol{x}} \text{ and}\\ \mathrm{Var}\left[\boldsymbol{x} \mid \boldsymbol{y}\right] &= \sigma^2\mathbf{N}^{-1}.\end{aligned}$$

This minimal MSE (conditional on $\boldsymbol{y}$) is

$$\begin{aligned}\mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right] &= \mathrm{E}\left[\|\hat{\boldsymbol{x}}-\boldsymbol{x}\|^2 \mid \boldsymbol{y}\right]\\ &= \hat{\boldsymbol{x}}^{\mathrm{t}}\hat{\boldsymbol{x}} - 2\hat{\boldsymbol{x}}\,\mathrm{E}\left[\boldsymbol{x} \mid \boldsymbol{y}\right] + \mathrm{E}\left[\boldsymbol{x}^{\mathrm{t}}\boldsymbol{x} \mid \boldsymbol{y}\right]\\ &= -\hat{\boldsymbol{x}}^{\mathrm{t}}\hat{\boldsymbol{x}} + \mathrm{Tr}\left[\mathrm{E}\left[\boldsymbol{x}\boldsymbol{x}^{\mathrm{t}} \mid \boldsymbol{y}\right]\right]\\ &= -\hat{\boldsymbol{x}}^{\mathrm{t}}\hat{\boldsymbol{x}} + \mathrm{Tr}\left[\sigma^2\mathbf{N}^{-1} + \hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^{\mathrm{t}}\right]\\ &= \mathrm{Tr}\left[\sigma^2\mathbf{N}^{-1}\right].\end{aligned}$$

3 Inexact MAP estimator

The following defines the inexact solution of the normal system (2) as an estimator and relates its MSE to that of the MAP estimator (further on called the *exact* MAP estimator).

3.1 Definition and properties

Let us define the *inexact MAP estimator* as an approximate solution to the normal system (2). Formally, the inexact estimator $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$ satisfies

$$\|\mathbf{A}^{\mathrm{t}}\boldsymbol{y} - \mathbf{N}\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\| = \tau.$$

Without knowledge of the process used to generate such an estimator, we will consider that all possible estimators $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$ with respect to $\boldsymbol{y}$ are equiprobable.

Lemma 1 *Consider the exact MAP estimator $\hat{\boldsymbol{x}}$ and fixed inexact MAP estimator $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$. We have*

$$\frac{\tau^2}{\|\mathbf{N}\|^2} \leq \mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] \leq \tau^2\|\mathbf{N}^{-1}\|^2.$$

Proof We develop

$$\begin{aligned}\mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] &= \mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \mathbf{N}^{-1}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right]\\ &= \|\mathbf{N}^{-1}(\mathbf{N}\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \mathbf{A}^{\mathrm{t}}\boldsymbol{y})\|^2.\end{aligned}$$

The upper bound is derived as

$$\mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] \leq (\|\mathbf{N}^{-1}\|\|\mathbf{N}\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \mathbf{A}^{\mathrm{t}}\boldsymbol{y}\|)^2 = \tau^2\|\mathbf{N}^{-1}\|^2,$$

by definition of the inexact estimator. The lower bound comes from

$$\mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] = \frac{\|\mathbf{N}\|^2}{\|\mathbf{N}\|^2}\|\mathbf{N}^{-1}(\mathbf{N}\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \mathbf{A}^{\mathrm{t}}\boldsymbol{y})\|^2 = \frac{\|\mathbf{N}\mathbf{N}^{-1}(\mathbf{N}\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \mathbf{A}^{\mathrm{t}}\boldsymbol{y})\|^2}{\|\mathbf{N}\|^2} = \frac{\tau^2}{\|\mathbf{N}\|^2},$$

which concludes the proof. □

Corollary 1 *Consider the exact MAP estimator $\hat{\boldsymbol{x}}$ and inexact MAP estimator $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$. We then have*

$$\frac{\tau^2}{\|\mathbf{N}\|^2} \leq \mathrm{MSE}\left[\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] - \mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right] \leq \tau^2\|\mathbf{N}^{-1}\|^2. \tag{9}$$

Proof Develop:

$$\begin{aligned}
\mathrm{MSE}\left[\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] &= \mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \boldsymbol{x}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] \\
&= \mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}} + \hat{\boldsymbol{x}} - \boldsymbol{x}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] \\
&= \mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] + 2\,\mathrm{E}\left[(\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}})^{\mathrm{t}}(\hat{\boldsymbol{x}} - \boldsymbol{x}) \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] + \mathrm{E}\left[\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \mid \boldsymbol{y}\right] \\
&= \mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] + 2(\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}})^{\mathrm{t}}(\hat{\boldsymbol{x}} - \mathrm{E}\left[\boldsymbol{x} \mid \boldsymbol{y}\right]) + \mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right].
\end{aligned}$$

Therefore, $\mathrm{MSE}\left[\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] - \mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right] = \mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right]$, so lemma 1 yields the inequalities (9). □

3.2 Stopping test for numerical methods

Corollary 1 suggests that the MSE (conditional on $\boldsymbol{y}$) of an inexact estimator is larger than that of the exact estimator and bounds the distance between the two. As we solve problem (1) through system (2) by a numerical method truncated once the norm of the residual reaches below a tolerance τ, we may set this tolerance in order to obtain a certain MSE. Formally, to get within a fraction α of the minimal MSE, one should set the tolerance so it satisfies

$$\tau^2\|\mathbf{N}^{-1}\|^2 \leq \alpha\,\mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right] = \alpha\sigma^2\,\mathrm{Tr}\left[\mathbf{N}^{-1}\right].$$

While σ^2 may be evaluated readily from the data, computing $\mathrm{Tr}\left[\mathbf{N}^{-1}\right]$ may be a costly endeavour. However, since the trace of a matrix corresponds to the sum of its eigenvalues, we have a lower bound for the MSE of the form

$$\mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right] = \sigma^2\,\mathrm{Tr}\left[\mathbf{N}^{-1}\right] \geq \sigma^2\left(\|\mathbf{N}^{-1}\| + \frac{n-1}{\|\mathbf{N}\|}\right) = \sigma^2\|\mathbf{N}^{-1}\|\left(1 + \frac{n-1}{\kappa\left(\mathbf{N}\right)}\right),$$

as the l_2 norm of a symmetric matrix corresponds to its eigenvalue of maximal absolute value and since matrix $\mathbf{N}$ is definite positive. Setting the stopping tolerance to

$$\tau = \sqrt{\frac{\alpha\sigma^2}{\|\mathbf{N}^{-1}\|}\left(1 + \frac{n-1}{\kappa\left(\mathbf{N}\right)}\right)}$$

then yields an inexact MAP estimator whose MSE is no more than α larger than the minimal MSE.

4 Extension to non-Gaussian priors

We will now extend the results of section 3 to the more general case of a non-Gaussian prior distribution. As this yields a space-variant stopping criterion, we will consider practical approaches to setting the tolerance.

4.1 MAP estimation with a non-Gaussian prior

The use of a quadratic penalization for the regularisation of the least-squares estimation problem (1) suits smooth objects best. However, any discontinuity that might be present in the object, such as edges between distinct regions in an image, will be significantly smoothed as estimated under such a penalization. Alternate penalty functions such as the generalized Gaussian Markov random field function (Bouman and Sauer, 1993) or the $l_2 l_1$ function (Charbonnier et al., 1997) achieve better preservation of discontinuities. The probabilistic framework that enables this alternative regularization strategy is the Gibbs prior distribution of the object.

The generalization of the estimation problem to a generic penalty function $\psi(u)$ is expressed as

$$\hat{\boldsymbol{x}} = \arg\min L(\boldsymbol{x} \mid \boldsymbol{y}) = \frac{1}{2}\|\boldsymbol{y} - \mathbf{A}\boldsymbol{x}\|^2 + \lambda R(\boldsymbol{x}), \tag{10}$$

where

$$R(\boldsymbol{x}) = \sum_{k=1}^{K} \nu_k \psi(\boldsymbol{\delta}_k^{\mathrm{t}} \boldsymbol{x}),$$

ν are penalization weight modifiers for each sub-term and $\boldsymbol{\delta}_k$ are the rows of matrix $\boldsymbol{\Delta}$, of dimensions $K \times n$. This matrix captures the features of the solution to penalize with respect to prior information known on the solution. For instance, in image reconstruction applications, it is typically composed by the concatenation of the identity and of finite difference operators, expecting the image to be composed of many black pixels and to be constant by regions.

The (exact) MAP estimator solves problem (10) by satisfying a stationarity condition over the objective function. Formally,

$$\nabla L(\hat{\boldsymbol{x}}) = \mathbf{A}^{\mathrm{t}}(\boldsymbol{y} - \mathbf{A}\hat{\boldsymbol{x}}) + \nabla R(\hat{\boldsymbol{x}}) = 0.$$

Such a stationary point is found using numerical procedures for nonlinear optimization, such as the Newton-Raphson method or a quasi-Newton iteration with a suitable globalization strategy. Larger-scale problems are often tackled with methods that necessitate neither the evaluation of Hessian of the objective nor any matrix approximation thereof, such as a nonlinear conjugate gradient generalization or a limited-memory quasi-Newton update. These methods are covered in depth by Nocedal and Wright (2000).

4.2 Linear extrapolation of the estimation problem

Whatever the numerical method, the final iterate is only approximately stationary. To determine the quality of inexact MAP estimators obtained as approximate solutions to problem (10), we must make some assumptions with regards to $\hat{\boldsymbol{x}}$. In the following, we restrict the analysis to penalty functions at least twice differentiable.

Hypothesis 1 *Consider $\hat{\boldsymbol{x}}$, the exact MAP estimator that solves problem* (10) *and $\boldsymbol{x_0} \in \mathbb{R}^n$ close to $\hat{\boldsymbol{x}}$. We then assume that, for $\boldsymbol{x}$ in a ball of radius $\|\boldsymbol{x_0} - \hat{\boldsymbol{x}}\|$ centered on $\hat{\boldsymbol{x}}$, we have*

$$L(\boldsymbol{x}) \approx \frac{1}{2}\|\boldsymbol{y} - \mathbf{A}\boldsymbol{x}\|^2 + \lambda \left(R(\boldsymbol{x_0}) + \nabla R(\boldsymbol{x_0})(\boldsymbol{x} - \boldsymbol{x_0}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x_0})^{\mathrm{t}} \nabla^2 R(\boldsymbol{x_0})(\boldsymbol{x} - \boldsymbol{x_0}) \right). \tag{11}$$

This hypothesis is justified by the log-likelihood function maximized as problem (10) being twice differentiable. It is well known that the behavior of sufficiently smooth function may be extrapolated faithfully in a small region around its stationary points by a second-order Taylor development (Nocedal and Wright, 2000). Where this extrapolation holds, we have

$$\begin{aligned} \nabla L(\boldsymbol{x}) &\approx -\mathbf{A}^{\mathrm{t}}(\boldsymbol{y} - \mathbf{A}\boldsymbol{x}) + \lambda \nabla R(\boldsymbol{x_0}) + \lambda \nabla^2 R(\boldsymbol{x_0})(\boldsymbol{x} - \boldsymbol{x_0}) \qquad (12) \\ &= \mathbf{A}^{\mathrm{t}}(\boldsymbol{y} - \mathbf{A}\boldsymbol{x}) + \lambda \sum_{k=1}^{K} \boldsymbol{\delta}_k \psi'(\boldsymbol{\delta}_k^{\mathrm{t}} \boldsymbol{x_0}) + \lambda \boldsymbol{\Delta}^{\mathrm{t}} \, \mathrm{D}_\psi \left[\boldsymbol{\Delta}\boldsymbol{x_0}\right] \boldsymbol{\Delta}(\boldsymbol{x} - \boldsymbol{x_0}), \end{aligned}$$

with $\mathrm{D}_\psi \left[\boldsymbol{u}\right]$, the *penalty curvature matrix*, a diagonal matrix of dimensions $K \times K$ for which element i of the diagonal is equal to $\psi''(\boldsymbol{u}_i)$. By defining

$$\begin{aligned} \mathbf{N_0} &= \mathbf{A}^{\mathrm{t}}\mathbf{A} + \lambda \boldsymbol{\Delta}^{\mathrm{t}} \, \mathrm{D}_\psi \left[\boldsymbol{\Delta}\boldsymbol{x_0}\right] \boldsymbol{\Delta} \text{ and} \\ \boldsymbol{z_0} &= \mathbf{A}^{\mathrm{t}}\boldsymbol{y} - \lambda[\nabla R(\boldsymbol{x_0}) - \nabla^2 R(\boldsymbol{x_0})\boldsymbol{x_0}], \end{aligned}$$

we re-express equation (12) as

$$\nabla L(\hat{\boldsymbol{x}}) \approx \mathbf{N_0}\hat{\boldsymbol{x}} - \boldsymbol{z_0} = 0.$$

and thus

$$\hat{\boldsymbol{x}} \approx \mathbf{N_0}^{-1}\boldsymbol{z_0},$$

which yields a closed form of the estimator from a reference point close to it.

Hypothesis 1 also allows us to extrapolate the distribution of $\boldsymbol{x}$ conditional on $\boldsymbol{y}$. As we have a single realization of $\boldsymbol{y}$, the likelihood of $\boldsymbol{x}$ conditional on $\boldsymbol{y}$ corresponds to the distribution. This implies that, from the extrapolated log-likelihood of equation (11), we can retrieve a Gaussian likelihood:

$$
\begin{aligned}
&\arg\min L(\boldsymbol{x}) \\
&= \arg\max\ -\frac{1}{2}(\boldsymbol{y}-\mathbf{A}\boldsymbol{x})^{\mathrm{t}}(\boldsymbol{y}-\mathbf{A}\boldsymbol{x}) \\
&\qquad -\lambda\left(R(\boldsymbol{x_0})+\nabla R(\boldsymbol{x_0})^{\mathrm{t}}(\boldsymbol{x}-\boldsymbol{x_0})+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x_0})^{\mathrm{t}}\nabla^2 R(\boldsymbol{x_0})(\boldsymbol{x}-\boldsymbol{x_0})\right) \\
&= \arg\max\ -\frac{1}{2}\left(\boldsymbol{y}^{\mathrm{t}}\boldsymbol{y}-\boldsymbol{y}^{\mathrm{t}}\mathbf{A}\boldsymbol{x}-\boldsymbol{x}^{\mathrm{t}}\mathbf{A}^{\mathrm{t}}\boldsymbol{y}+\boldsymbol{x}^{\mathrm{t}}\mathbf{A}^{\mathrm{t}}\mathbf{A}\boldsymbol{x}\right) \\
&\qquad -\lambda\left(\frac{1}{2}\nabla R(\boldsymbol{x_0})^{\mathrm{t}}\boldsymbol{x}+\frac{1}{2}\boldsymbol{x}^{\mathrm{t}}\nabla R(\boldsymbol{x_0})-\nabla R(\boldsymbol{x_0})^{\mathrm{t}}\boldsymbol{x_0}\right) \\
&\qquad -\frac{\lambda}{2}\left(\boldsymbol{x}^{\mathrm{t}}\nabla^2 R(\boldsymbol{x_0})-\boldsymbol{x}^{\mathrm{t}}\nabla^2 R(\boldsymbol{x_0})\boldsymbol{x_0}-\boldsymbol{x_0}^{\mathrm{t}}\nabla^2 R(\boldsymbol{x_0})\boldsymbol{x}+\boldsymbol{x_0}^{\mathrm{t}}\nabla^2 R(\boldsymbol{x_0})\boldsymbol{x_0}\right) \\
&= \arg\max\ -\frac{1}{2}\boldsymbol{x}^{\mathrm{t}}(\mathbf{A}^{\mathrm{t}}\mathbf{A}+\lambda\nabla^2 R(\boldsymbol{x_0}))\boldsymbol{x} \\
&\qquad +\frac{1}{2}\boldsymbol{x}^{\mathrm{t}}\left(\mathbf{A}^{\mathrm{t}}\boldsymbol{y}+\lambda\left(\nabla R(\boldsymbol{x_0})-\nabla^2 R(\boldsymbol{x_0})\boldsymbol{x_0}\right)\right) \\
&\qquad +\frac{1}{2}\left(\mathbf{A}^{\mathrm{t}}\boldsymbol{y}+\lambda\left(\nabla R(\boldsymbol{x_0})-\nabla^2 R(\boldsymbol{x_0})\boldsymbol{x_0}\right)\right)^{\mathrm{t}}\boldsymbol{x} \\
&= \arg\max -\frac{1}{2\sigma^2}\left(\boldsymbol{x}^{\mathrm{t}}\mathbf{N_0}\boldsymbol{x}-\boldsymbol{x}^{\mathrm{t}}\mathbf{N_0}\mathbf{N_0}^{-1}\boldsymbol{z_0}-\boldsymbol{z_0}^{\mathrm{t}}\mathbf{N_0}^{-1}\mathbf{N_0}\boldsymbol{x}+\boldsymbol{z_0}^{\mathrm{t}}\mathbf{N_0}^{-1}\mathbf{N_0}\mathbf{N_0}^{-1}\boldsymbol{z_0}\right) \\
&= \arg\max \log\exp\left(-\frac{(\boldsymbol{x}-\mathbf{N_0}^{-1}\boldsymbol{z_0})^{\mathrm{t}}\mathbf{N_0}(\boldsymbol{x}-\mathbf{N_0}^{-1}\boldsymbol{z_0})}{2\sigma^2}\right)-\log\left((2\pi)^{n/2}\left|\sigma^2\mathbf{N_0}^{-1}\right|\right) \\
&= \arg\max \frac{1}{(2\pi)^{n/2}\left|\sigma^2\mathbf{N_0}^{-1}\right|}\exp\left(-\frac{(\boldsymbol{x}-\mathbf{N_0}^{-1}\boldsymbol{z_0})^{\mathrm{t}}\mathbf{N_0}(\boldsymbol{x}-\mathbf{N_0}^{-1}\boldsymbol{z_0})}{2\sigma^2}\right)\approx f_{\boldsymbol{x}}(\boldsymbol{x}\mid\boldsymbol{y}),
\end{aligned}
\tag{13}
$$

the last objective yielding an extrapolated form of the probability density of $\boldsymbol{x}$ conditional on $\boldsymbol{y}$. This suggests that the extrapolated exact MAP estimator

$$\hat{\boldsymbol{x}} = \mathbf{N_0}^{-1}\boldsymbol{z_0} = \mathrm{E}\left[\boldsymbol{x}\mid\boldsymbol{y}\right]$$

minimizes the MSE conditional on $\boldsymbol{y}$. From the extrapolated form of the density of $\boldsymbol{x}$ conditional on $\boldsymbol{y}$(equation (13)), this MSE is expressed as

$$\mathrm{MSE}\left[\boldsymbol{x}\mid\boldsymbol{y}\right] = \sigma^2\,\mathrm{Tr}\left[\mathbf{N_0}^{-1}\right].$$

Definition 1 *The* inexact MAP estimator $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$*in the context of a nonquadratic penalization satisfies*

$$\|\nabla L(\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}})\| = \tau.$$

With respect to hypothesis 1, this suggests that for reference point $\boldsymbol{x_0}$ close enough to $\hat{\boldsymbol{x}}$ and $\tau \leq \|\boldsymbol{x_0}-\hat{\boldsymbol{x}}\|$, the inexact estimator satisfies

$$\|\boldsymbol{z_0}-\mathbf{N_0}\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\| = \tau.$$

Moreover, for τ small enough, $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$ itself may be used as "reference point" for the extrapolation. This completes the set of arguments that generalize the results of section 3 to estimators developed with a non-Gaussian prior distribution, in a small neighborhood of $\hat{\boldsymbol{x}}$. The following lemma and corollary are given without explicit proof, as the arguments are very close to those in the Gaussian prior case, with $\mathbf{N}$ and $\mathbf{A}^t\boldsymbol{y}$ replaced respectively with $\mathbf{N_0}$ and $\boldsymbol{z_0}$.

Lemma 2 *Let $\hat{\boldsymbol{x}}$ and $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$ be the exact and inexact MAP estimators, respectively. For small τ, we have*

$$\frac{\tau^2}{\|\mathbf{N_0}\|^2} \leq \mathrm{E}\left[\|\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} - \hat{\boldsymbol{x}}\|^2 \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] \leq \tau^2\|{\mathbf{N_0}}^{-1}\|^2.$$

Corollary 2 *Let $\hat{\boldsymbol{x}}$ and $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$ be the exact and inexact MAP estimators, respectively. For small τ, we have*

$$\frac{\tau^2}{\|\mathbf{N_0}\|^2} \leq \mathrm{MSE}\left[\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}} \mid \boldsymbol{y}, \tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}\right] - \mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right] \leq \tau^2\|{\mathbf{N_0}}^{-1}\|^2.$$

4.3 Stopping test revisited

Corollary 2 may be interpreted as a stopping test for the numerical method used to solve problem (10). This test involves the Euclidean norm of the extrapolated normal system (12), which more generally corresponds to the norm of the gradient of the objective, a common measure of stationarity of the solution in nonlinear programming. To generate an inexact estimator for which the MSE is decreased to within a fraction α of the minimal MSE, one should stop at iterate $\boldsymbol{x}_k$ such that

$$\|\nabla L(\boldsymbol{x}_k)\| \leq \tau = \sqrt{\frac{\alpha\,\mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right]}{\|{\mathbf{N_0}}^{-1}\|^2}}.$$

Once again, $\mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right]$ is hard to evaluate exactly, as matrix ${\mathbf{N_0}}^{-1}$ is not represented explicitly. However, a lower bound is determined as

$$\mathrm{MSE}\left[\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right] \geq \sigma^2\|{\mathbf{N_0}}^{-1}\|\left(1 + \frac{n-1}{\kappa\left(\mathbf{N_0}\right)}\right),$$

which may be evaluated as $\|{\mathbf{N_0}}^{-1}\|$ corresponds to the inverse of the smallest eigenvalue of $\mathbf{N_0}$. The stopping tolerance may then be rewritten as

$$\tau = \sqrt{\frac{\alpha\sigma^2}{\|{\mathbf{N_0}}^{-1}\|}\left(1 + \frac{n-1}{\kappa\left(\mathbf{N_0}\right)}\right)}. \tag{14}$$

It appears from relation (14) that the tolerance selection depends on the reference point from which the log-likelihood is extrapolated. This entails that the tolerance should not be computed from any $\boldsymbol{x_0}$, as the penalty curvature matrix $\mathrm{D}_\psi\left[\boldsymbol{\Delta x_0}\right]$ has a significant influence on $\|{\mathbf{N_0}}^{-1}\|$. For instance, edge-preserving penalty functions (such as the GGMRF and l_2l_1

functions mentioned earlier) have maximal curvature at $u = 0$, making the smallest eigenvalue of matrix $\nabla^2 R(0)$ rather large. Therefore, a stopping tolerance computed from this eigenvalue will be larger than it should to reach the expected MSE.

Therefore, the actual tolerance should be computed from some $\boldsymbol{x_0}$ that resembles $\hat{\boldsymbol{x}}$ to a certain degree, in order for the curvature matrix to be close to that at the exact solution. One could then generate an inexact estimator $\tilde{\boldsymbol{x}}_{\boldsymbol{\tau}}$ for a value of τ that is small but does not entail too much computations. The final stopping tolerance would be computed from this pilot estimator. If it exceeds the initial tolerance, the pilot estimator satisfies the MSE constraint; otherwise, the nonlinear solver would be restarted from the pilot value and made to reach stationarity to the final tolerance. Another interesting approach is to compute the reference point by a fast approximation procedure when such a procedure exists for the application considered. For example, in context of transmission tomography reconstruction, one may obtain the reference point by the filtered backprojection algorithm (Kak and Slaney, 1987).

5 Concluding remarks

We have exposed a relation between the defining feature of an inexact estimator and its conditional MSE. From this, we have formulated a stopping criterion for numerical methods for solving the estimation problem, based on the stationarity measure of the iterate.

Besides the cavalier treatment of the prior distribution of $\boldsymbol{x}$ under hypothesis 1, the salient weakness of the development above is the systematic undervaluation of the MSE of the exact estimator. This yields unduly low stopping tolerances, leading to the execution of more iterations than necessary to ascertain the statistic properties of the solution we accept. Yet, for an application of edge-preserving tomographic reconstruction, this exact framework has proposed sensible tolerances, for which reconstruction runtimes were acceptable with respect to data and image dimensions.

This said, for applications such as deconvolution in signal and image processing, the trace of the covariance matrix of the exact estimator can be more easily computed. This should entail looser converge criteria. We will investigate this issue in depth in an expanded version of this paper, where the general Gaussian measure model with Gaussian prior will be fully developed.

References

Charles A. Bouman and Ken D. Sauer. A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Transactions on Image Processing*, IP-2(3):296–310, July 1993.

Pierre Charbonnier, Laure Blanc-Féraud, Gilles Aubert, and Michel Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, IP-6(2):298–311, February 1997.

Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publishers, 1996.

Jérôme Idier, editor. *Approche Bayésienne pour les Problèmes Inverses.* Traitement du Signal et de l'Image. Hermes Science – Lavoisier, Paris, France, 2001.

Avinash C. Kak and Malcolm Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, New York, NY, 1987.

C. Tim Kelley. *Iterative Methods for Linear and Nonlinear Equations.* SIAM, Philadelphia, 1995.

Jorge Nocedal and Stephen J. Wright. *Numerical Optimization.* Operations Research. Springer Verlag, New York, NY, 2000.

K. Sam Shanmugan and Arthur M. Breipohl. *Random Signals: Detection, Estimation and Data Analysis.* John Wiley & Sons, 1988.