Regularization networks for inverse problems: A state-space approach

Giuseppe De Nicolao a,1 Giancarlo Ferrari-Trecate a,b

a Dipartimento di Informatica e Sistemistica, Università degli Studi di Pavia, Via Ferrata 1, 27100 Pavia, Italy Phone: +39.382.505484, Fax: +39.382.505373 E-mail: Giuseppe.Denicolao@unipv.it

b Currently: INRIA, Domaine de Voluceau, Rocquencourt B.P.105, 78153 Le Chesnay Cedex, France +33.1.39635327, Fax: +33.1.39635786 E-mail: Giancarlo.Ferrari-Trecate@inria.fr

Abstract

Linear inverse problems with discrete data are equivalent to the estimation of the continuous-time input of a linear dynamical system from samples of its output. The solution obtained by means of regularization theory has the structure of a neural network similar to classical RBF networks. However, the basis functions depend in a nontrivial way on the specific linear operator to be inverted and the adopted regularization strategy. By resorting to the Bayesian interpretation of regularization, we show that such networks can be implemented rigorously and efficiently whenever the linear operator admits a state-space representation. An analytic expression is provided for the basis functions as well as for the entries of the matrix of the linear system used to compute the weights. The results are illustrated through a deconvolution problem where the spontaneous secretory rate of Luteinizing Hormone (LH) of the hypophisis is reconstructed from measurements of plasma LH concentrations.

Key words: Inverse problems, neural networks, regularization, Bayesian estimation, deconvolution, Kalman filter.

1 Introduction

Inverse problems are almost ubiquitous in engineering and science. To cite some examples, they play a crucial role in signal equalization (Douglas et al. 1999), echo cancellation (Muller and Elmirghani 1999), image restoration (Tong 1995), system identification (Yellin and Friedlander 1999, Magni et al. 1998), physiological systems (De Nicolao et al. 1997), physiology and analysis of DNA segmentation (Ives et al. 1994). These problems are “inverse”, because reconstructing the input (the cause) of a system from output measurements (the effects), amounts to reversing of the causality axis.

A classical approach to the solution of inverse problems is the regularization method of Tikhonov and Arsenin (1977), see also (Bertero 1989) for an extensive survey. The unknown input is found by minimizing a cost functional consisting of the sum of two terms: the first is the usual least squares fit and the second is a stabilizer that penalizes the estimates that are too “irregular”. For instance, the stabilizer term is often chosen as the integral of the squared first or second derivative of the estimated input signal, although other choices are possible.

Rather interestingly, also the problem of reconstructing a real function of several variables from a finite number $N$ of noisy measurements is a particular type of inverse problem. Girosi et al. (1995) have shown that in this case the regularized solution is a neural network, called regularization network, whose output is the linear combination, through proper weights, of $N$ basis functions. The weights are the solution of a linear system and their computation requires $O(N^3)$ operations. The basis functions are the Green’s functions of the Gram’s operator associated with the stabilizer. If the stabilizer exhibits radial symmetry, the basis functions are radially symmetric as well and an RBF (Radial Basis Function) net-
work is obtained. In general, it is not trivial to compute the Green's functions. However, in some specific case, they can be explicitly evaluated, as happens for polynomial and thin-plate splines or for the Gaussian stabilizer (Giroi et al. 1995).

Coming back to generic linear inverse problems, it is relatively straightforward to show that the solution provided by the Tychonov's method has still the structure of a regularization network. Once the basis functions are constructed, the regularization network coincides with the Bayesian estimator obtained by assuming that the unknown function is a Gaussian stochastic process with suitable autocovariance. Then, under suitable assumptions on the stabilizer, the regularization network remains the same.

However, the use of regularization networks for the solution of inverse problems suffers from two main drawbacks. A first problem is to obtain a closed-form expression of the basis functions because they depend also on the linear operator to be inverted. For instance, even if the stabilizer is radially symmetric, the basis functions, in general, will not be. A second problem is that the weight calculation requires \(O(N^3)\) operations thus rendering the algorithm unsuitable for large data sets.

The aim of the present paper is to show that regularization networks can be efficiently implemented for an important class of inverse problems where both the input and the output are scalar functions of a real variable. For this purpose, the Bayesian interpretation of regularization networks is exploited (Section 3). In fact, under suitable assumptions on the stabilizer, the regularization network coincides with the Bayesian estimator obtained by assuming that the unknown function is a Gaussian stochastic process with suitable autocovariance. Then, the basis functions are the autocovariance of the output of the linear operator fed by the Gaussian process. The class of problems we consider is characterized by the fact that both the linear operator to be inverted and the stochastic process describing the unknown function admit a state-space representation (Section 4). Under this assumption, the basis functions will be characterized and computed explicitly in Section 5.

For what concerns the calculation of the weights, the Bayesian interpretation shows that it amounts to the inversion of a covariance matrix. By exploiting a recent result (De Nicolao and Ferrari-Trecate 2001), this can be done in \(O(N)\) operations through a state-space algorithm based on Kalman filtering.

To illustrate the results, in Section 6 we consider the deconvolution of a time-series of Luteinizing Hormone (LH) plasma concentrations. This inverse problem arises in the estimation of the spontaneous LH secretion rate of the hypophysis and is a prototype for the estimation of the secretion rate of other pituitary hormones.

2 Problem statement

Consider the linear inverse problem with discrete data given by

\[
y_i = \int_0^{t_i} h(t_i, \tau)f(\tau)d\tau + v_i, \quad i = 1, 2, \ldots, N,
\]

where \(f(\cdot) : \mathbb{R} \rightarrow \mathbb{R}\) is the unknown signal to be reconstructed, \(h(\cdot, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}\) is the impulse response characterizing the linear operator which is assumed to be causal, i.e., \(h(t, \tau) = 0, \forall t < \tau\) and \(v_i\) is the measurement noise associated with the \(i\)-th measurement \(y_i\) taken at time \(t_i\). We assume that the time samples \(t_i\) are ordered (i.e., \(t_1 < t_2 < \ldots < t_N\)) and that both \(f(\cdot)\) and \(h(\cdot, \cdot)\) are squared integrable functions. Moreover, the measurement errors \(v_i\) are uncorrelated Gaussian zero-mean random vectors with \(\text{Var}[v_i] = \sigma_i^2\). By defining

\[
H_if = \int_0^{t_i} h(t_i, \tau)f(\tau)d\tau,
\]

the integral equation (1) can be written in the operatorial form

\[
y_i = H_if + v_i, \quad i = 1, 2, \ldots, N.
\]

It is well known that the inverse problem (2) is ill-posed unless some prior knowledge about the function \(f\) is introduced (Bertero et al. 1985). A popular approach is the so-called regularization method (Tikhonov and Arsenin 1977), (Bertero 1989) which looks for the function \(\hat{f}(\cdot)\) that minimizes the functional

\[
J[f] = \sum_{i=1}^{N} \frac{(y_i - H_if)^2}{\sigma_i^2} + \gamma \|Pf\|^2,
\]

where \(P\) is a penalty operator, \(\|\cdot\|^2\) is the norm on the space to whom \(f\) belongs, and the positive real number \(\gamma > 0\) is the so-called regularization parameter. The operator \(P\) is chosen to reflect the prior knowledge and typically is such that \(\|Pf\|^2\) is large whenever \(f\) is "irregular", i.e. wiggly. Then, it is apparent that

\[
\hat{f} = \arg\min_f J[f]
\]
is the result of a compromise between the least squares fit (measured by the first term in the RHS of (3)) and the smoothness of $f$ (measured by $\|Pf\|^2$). The balance between the two terms is controlled by the parameter $\gamma$, that can be tuned by means of various techniques, including Cross Validation, Generalized Cross Validation (Wahba 1990) and Maximum Likelihood (MacKay 1992, De Nicolao et al. 1997).

**Remark 1.** Girosi et al. (1995) considered the regression problem

$$y_i = f(t_i) + v_i, \quad i = 1, 2, \ldots, N,$$

which is just equivalent to (1) with $h(t, \tau) = \delta(t - \tau)$, where $\delta(t)$ is the Dirac’s delta. In this special case, the regularized estimate $\hat{f}(t)$ has the structure of an RBF-like network which they named regularization network. More precisely,

$$\hat{f}(t) = \sum_{i=1}^{N} c_i G(t, t_i)$$

where the scalar coefficients $c_i$ can be computed by solving a linear system. The activation functions of the neurons are given by $G(t, t_i)$, where $G(t, s)$ is the Green’s function of the self-adjoint operator $P^*P$, i.e.

$$P^*PG(t, s) = \delta(t - s).$$

Under the assumption that $\|Pf\|^2$ is rotationally and translationally invariant, the Green function has radial symmetry, that is $G(t, s) = G(\|t - s\|)$, so that (5) is an RBF network in strict sense.

Coming back to the more general inverse problem (1) it is not surprising that the regularized solution $\hat{f}(t)$, obtained from the minimization of (3), has again a network structure.

**Proposition 1 (Wahba (1977)).** Letting $c = [c_1 \ c_2 \ \cdots \ c_N]^T$, $y = [y_1 \ y_2 \ \cdots \ y_N]^T$, the minimizer of (3) is given by

$$\hat{f}(t) = \sum_{i=1}^{N} c_i \eta_i(t) = \sum_{i=1}^{N} c_i G(t, t_i),$$

$$\eta_i(t) = \int_{0}^{\infty} h(t, \tau) G(t, \tau) d\tau,$$

where $G(t, s)$ satisfies (6) and $c$ is such that

$$(Q + \gamma D)c = y, \quad (Q)_{ij} = Q(t_i, t_j),$$

$\gamma = \frac{\lambda^2}{\lambda_0^2}$

$D = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix}$

$$Q(t, s) = \int_{0}^{\infty} \int_{0}^{\infty} h(t, \tau) G(t, \xi) h(\xi, s) d\tau d\xi$$

The network (7)-(10) is a natural extension of the regularization network model introduced by Girosi et al. (1995). In the inverse problem considered herein, differently form the regression problem studied by Girosi et al. (1995), the activation functions $\eta_i(t)$ do not depend only on the penalty operator $P$, but also on the kernel $h(t, \tau)$, see (8). As a consequence, even when $\|Pf\|^2$ is rotationally invariant, the activation functions, in general, will not. Moreover, even when the penalty operator has a simple structure, an analytic characterization of the functions $\eta_i(t)$ is not trivial and one may have to numerically evaluate the integrals (8) and (10).

The main aim of the present paper is to provide an algorithm for the computation of the activation functions $\eta_i(t)$ that holds for linear inverse problems admitting a state-space representation. This method avoids any form of numerical integration and in some significant cases leads to closed-form expressions. Another important issue is the calculation of the weights $c_i$ which usually requires $O(N^3)$ operations (that is the computational complexity of solving the linear system (9)). Remarkably, for the class of problems considered in the present paper, the complexity can be reduced to $O(N)$ by exploiting the algorithm based on Kalman filtering proposed by De Nicolao and Ferrari-Trecate (2001).

3 The Bayesian interpretation of regularization

The derivation of the main results relies on the Bayesian interpretation of regularization which is now briefly summarized. In Bayesian estimation the prior knowledge on the unknown function $f$ is embodied in the so-called prior probability function $p(f)$. Once the vector of observations $y$ is available, the Bayes rule can be used to compute the posterior probability function as

$$p(f|y) \propto p(f)p(y|f)$$

where $p(y|f)$ is the likelihood of the data. The Bayes estimate $f^B$ of the unknown function is then given by the posterior expectation, i.e. $f^B = E[f|y]$.

Hereafter, it is assumed that the unknown function $f$ is modeled as a Gaussian stochastic process independent of the errors $v_i$ and with prior probability distribution

$$p(f) \propto e^{-\frac{1}{2\lambda^2} \langle f, Pf \rangle}$$

where $\langle \cdot, \cdot \rangle$ is a scalar product in the space to whom $f$ belongs, $\lambda^2$ is a positive scale factor and $P$ is the penalty operator introduced previously. In other words, $f$ is a zero-mean Gaussian process with covariance function $E[f(t)f(s)]$ proportional to the Green’s function $G(t, s)$.

**Example 1.** Assume that $f(0) = 0$ and $Pf \equiv f(\tau)$. Then, $\lambda^2 G(t, s)$ turns out to be the autocovariance function of the stochastic process $f$ satisfying

$$df(t) = dw(t), \quad f(0) = 0$$
where \( w(t) \) is a Wiener process with incremental variance \( \lambda^2 dt \). It is immediate to see that \( P \) is the whitening filter of the stochastic process \( f \) defined in (12). This holds true also in general so that the process \( f \) with autocovariance \( G(t,s) \) is just the process whose whitening filter is equal to the operator \( P \).

**Proposition 2 (Girosi et al. (1995)).** If \( \lambda^2 = 1/\gamma \) then \( B \) is the minimizer of (3), that is \( B = \hat{f} \).

**Proof.** By exploiting Gaussianity, the posterior mean \( f \) can be written according to the well known formula, see e.g. Anderson and Moore (1979),

\[
    f = E[f|y] = E[f] Cov[y]^{-1}y, \tag{13}
\]

where \( z = [z_1 z_2 \ldots z_N]^T \) is the vector of the uncorrupted observations, i.e. \( z_i = h_i f \). In view of (1), (2) and (7)-(10) it is not difficult to see that

\[
    \eta(t) = E[f(t)z_i] \tag{14}
\]

\[
    Q + D = Cov[y]. \tag{15}
\]

Then, the thesis directly follows by comparing (13) and (7).

\[\square\]

### 4 State-space representation of the prior and the linear operator

In this Section, inverse problems admitting a state-space representation are considered. More precisely, it is assumed that the Gaussian Process \( f \) is the output of the single-input single-output linear time-invariant system

\[
    \Sigma : \begin{cases} 
        dx(t) = \tilde{A} \tilde{x}(t) dt + \tilde{B} \tilde{w}(t) \\
        f(t) = C \tilde{x}(t), \quad \tilde{x}(t_0) = 0
    \end{cases} \tag{16}
\]

where \( \tilde{x}(t) \in \mathbb{R}^l \) is the state, \( \tilde{w}(t) \in \mathbb{R} \) is a Wiener process with incremental variance \( \lambda^2 dt \), and \( \tilde{A}, \tilde{B}, \tilde{C} \) are matrices of suitable dimensions. In other words, \( f \) is a non-stationary and Markovian Gaussian process (Anderson and Moore 1979). Note that we have introduced an initial time \( t_0 \) where the initial state is zero. This is not particularly restrictive because, as demonstrated in the numerical example (Section 6), one can take \( t_0 \ll t_1 \) in order to allow for high uncertainty on \( x(t_1) \). As an alternative, one could regard \( \tilde{x}(t_0) \) as a Gaussian random vector whose prior uncertainty is assigned through the covariance matrix \( \text{Cov}[\tilde{x}(t_0)] \).

**Remark 2.** As a particular case, the model (16) covers the case of polynomial smoothing splines. In fact, with the choice

\[
    \tilde{A} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \tilde{C} = [0 \ldots 0 1],
\]

the output \( f(t) \) of (16) is an \((l-1)\)-fold integrated Wiener process. Then, in the function reconstruction case (e.g. \( h(t,\tau) = \delta(t-\tau) \), see (4)) the regularized (and Bayes) estimate is a \( (2l-1)\)-th order smoothing spline with zero boundary conditions (Wahba 1990).

Concerning the impulse response \( h(t,\tau) \), for the sake of simplicity we assume that it satisfies the stationarity condition \( h(t,\tau) = h(t-\tau) \). Moreover it is assumed that \( h(t,\tau) \) is the impulse response of the single-input single-output linear system

\[
    \Sigma : \begin{cases} 
        d\tilde{x}(t) = \tilde{A} \tilde{x}(t) dt + \tilde{B} \tilde{f}(t) dt \\
        z(t) = \tilde{C} \tilde{x}(t) \\
        \tilde{x}(t_0) = 0
    \end{cases}, \quad \tilde{x}(t) \in \mathbb{R}^r \tag{17}
\]

for suitable choices of the \( \tilde{A}, \tilde{B} \) and \( \tilde{C} \) matrices. It is apparent that the stationarity assumption could be removed by considering time varying matrices \( \tilde{A}(t), \tilde{B}(t) \) and \( \tilde{C}(t) \).

From (16) and (17) it follows that the overall model is

\[
    \Sigma : \begin{cases} 
        dx(t) = Ax(t) dt + Bdw(t) \\
        z(t) = Cx(t) \\
        x(t_0) = x_0 \\
        y_i = z(t_i) + v_i, \quad i = 1,2,\ldots,N \tag{18}
    \end{cases}
\]

\[
    f(t) = Lx(t) \tag{19}
\]

where

\[
    A = \begin{bmatrix} \tilde{A} & 0 \\ \tilde{B} C & \tilde{A} \end{bmatrix}, \quad B = \begin{bmatrix} \tilde{B} \\ 0 \end{bmatrix}, \quad x(t) = \begin{bmatrix} \tilde{x}(t) \\ \tilde{x}(t) \end{bmatrix},
\]

\[
    x_0 = \begin{bmatrix} \tilde{x}(t_0) \\ \tilde{x}(t_0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},
\]

\[
    C = \begin{bmatrix} 0 & \tilde{C} \end{bmatrix}, \quad L = \begin{bmatrix} \tilde{C} & 0 \end{bmatrix}
\]

The block diagram summarizes the deconvolution problem is reported in Figure 1.
The state-space representation of Section 4 is now exploited in order to derive an algorithm for the computation of the basis function $\eta_i(t)$ and the entries of the matrix $Q$ defined in (9).

**Theorem 3.** Let $X(t) = \mathbb{E} \left[ x(t)x(t)' \right]$. Then

$$\eta_k(t) = \begin{cases} Le^{A(t-t_k)}X(t_k)C' & t_k < t, \\
LX(t_k)e^{A'(t_k-t)}C' & t \leq t_k, \end{cases} \tag{21}$$

$$E \left[ z(t_i)z(t_j)' \right] = \begin{cases} Ce^{A(t_j-t_i)}X(t_i)C' & t_i < t_j, \\
CX(t_i)e^{A'(t_i-t_j)}C' & t_j \leq t_i, \end{cases} \tag{22}$$

$$X(t) = \lambda^2 \int_0^{t-t_0} e^{A\tau}BB'e^{A\tau}d\tau, \tag{23}$$

**Proof.** It is well known that the state covariance matrix $X(t), t \geq t_0$, is the solution of the Lyapunov differential equation

$$\dot{X}(t) = AX(t) + X(t)A' + \lambda^2 BB' \tag{24}$$

with initial condition $X(t_0) = 0_{(t+r)x(t+r)}$. The explicit solution of (24) is given by (23). It is also immediate to see that

$$E \left[ x(t)x(\tau)' \right] = \begin{cases} e^{A(t-\tau)}X(\tau) & \tau < t, \\
X(\tau)e^{A'(\tau-t)} & t \leq \tau, \end{cases} \tag{25}$$

As for the basis functions, from (14), (18) and (20) we obtain, for $k = 1, 2, \ldots, N$,

$$\eta_k(t) = \mathbb{E}[f(t)x_k] = LE \left[ x(t)x(t_k)' \right] C',$$

and, by using (25), one derives (21). Finally, from (18), it is apparent that

$$E \left[ z(t_i)z(t_j)' \right] = CE \left[ x(t_i)x(t_j)' \right] C',$$

and, using again (25), expression (22) is proved. \hfill \Box

**Proposition 4.** The entries $(Q)_{ij}$ of the matrix $Q$ satisfy

$$(Q)_{ij} = \frac{E[z(t_i)z(t_j)']}{\lambda^2}. \tag{26}$$

**Proof.** Since $f$ and $v_i$ are statistically independent, one obtains

$$\text{Cov}[y] = \text{Cov} \left[ \begin{bmatrix} H_1f & \cdots & H_Nf \end{bmatrix}' \right] + D = \text{Cov} \left[ \begin{bmatrix} z(t_1) & \cdots & z(t_N) \end{bmatrix}' \right] + D.$$

Then, by direct comparison with (15), we have $Q = \text{Cov} \left[ \begin{bmatrix} z(t_1) & \cdots & z(t_N) \end{bmatrix}' \right]$. By recalling that $\gamma = \frac{1}{\lambda^2}$ and $\mathbb{E}[z] = 0$, the thesis follows. \hfill \Box

Formulas (21), (22), (23) and (26) show that one can compute explicitly both $\eta_k(t)$ and $Q$ if closed-form expressions of $X(t)$ and $e^{At}$ are available. In any case, $X(t)$ can be evaluated without resorting to numerical integration by using the following formula based on the matrix exponential (Van Loan 1978)

$$\tilde{F} = \begin{bmatrix} -A & BB' \\
0 & A \end{bmatrix}, \quad \begin{bmatrix} \Phi_1 & \Psi \\
0 & \Phi_2 \end{bmatrix} = e^{\tilde{F}(t-t_0)},$$

$$\int_0^{t-t_0} e^{A\tau}BB'e^{A\tau}d\tau = \Phi_2^t \Psi.$$

where $\Phi_1 \in \mathbb{R}^{(t+r)x(t+r)}$. In the remaining part of this Section, we specialize the results of Theorem 3 to deconvolution problems where the impulse response has the structure

$$h(t) = \sum_{i=1}^r a_i e^{-\alpha_i t}, \quad r \in \mathbb{N}_0$$

By the way, impulse responses of this type find an important application in the compartmental modeling of physiological systems (see the example in Section 6). Then, such $h(t)$ admits the state-space representation (17) with

$$\tilde{A} = \begin{bmatrix} -\alpha_1 & 0 & \cdots & 0 \\
0 & -\alpha_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & -\alpha_r \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 1 \\
1 \\
\vdots \\
1 \end{bmatrix},$$

$$\tilde{C} = \begin{bmatrix} a_1 & a_2 & \ldots & a_r \end{bmatrix}.$$
\( \hat{A} = 0, \ B = \hat{C} = 1. \) Then, the matrices \( A, B, \) and \( C \) of system (18) are

\[
A = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
1 & -\alpha_1 & 0 & \cdots & 0 \\
1 & 0 & -\alpha_2 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & 0 & \cdots & -\alpha_r \\
\end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ \underline{0_r} \end{bmatrix}, \quad (27)
\]

(28)

Lemma 5. Consider the state-space model (27)-(28). Then, it holds that

\[
\eta_k(t) = \begin{cases} 
  \frac{\eta_k(t_k)}{\chi'} & t > t_k, \\
  (t-t_0)\hat{G}(t_k-t)^t \hat{C} + s(t)^t e^{A(t_k-t)} \hat{C}^t & t_0 \leq t \leq t_k, \\
\end{cases} \quad (29)
\]

\[
E[z(t_i)z(t_j)] = \begin{cases} 
  \alpha^2 \hat{G}[\eta_i(t_j) + s(t_j)^t e^{A(t_j-t)} s(t_j)] \hat{C} & t_i \geq t_j, \\
  E[z(t_i)z(t_j)] & t_i < t_j, \\
\end{cases} \quad (30)
\]

(31)

(32)

(33)

(34)

(35)

(36)

Proof. Due to the special structure of system (27), the matrices \( e^{At} \) and \( X(t) \) can be explicitly computed as in (31)-(36). Then, Theorem 3 can be applied in order to compute \( \eta_k(t) \) and \( E[z(t_i)z(t_j)] \).

In order to completely solve the inverse problem, the last step is to compute the weight vector \( c \) through the solution of the linear system (9). In general, the computational complexity of this operation is \( O(N^3) \). However, from (13) it is apparent that the problem amounts to inverting the covariance matrix of the observations. Recently, De Nicolao and Ferrari-Trecate (2001) have shown that if the observed process is Markovian, the linear system (9) can be solved in \( O(N) \) operations by Kalman filtering techniques.

6 Deconvolution of LH data

The algorithm worked out in the previous sections has been tested on a time-series of plasma LH (luteinizing hormone) concentrations in a normal subject consisting of the samples \( y_k, k = 1, \ldots, 40 \), collected with a uniform 5-min. sampling period \( t_k = 5(k-1) \) (Genazzani et al. 1990). The aim is to reconstruct the Instantaneous Secretion Rate (ISR), i.e. the flux of hormone from the hypothysis into circulation. To ease the comparison with a previous discretization approach, we analyze the same set of data presented in De Nicolao et al. (1997). The deconvolution of LH data with 5-min. sampling is an adequate benchmark problem due to its ill-conditioning caused by the relatively long time constants of hormone decay in the circulation (De Nicolao and Liberati 1993).

The impulse response describing the hormone decay in the circulation is given by the second-order model (\( r = 2 \)):

\[
\hat{g}(t) = a_1 e^{-\alpha_1 t} + a_2 e^{-\alpha_2 t}
\]

where \( a_1, a_2, \alpha_1, \alpha_2 \), are average values estimated in a population of normal subjects \( (a_1 = 0.615 \text{ mIU min}^{-1}, a_2 = 0.385 \text{ mIU min}^{-1}, \alpha_1 = 3.87 \times 10^{-2} \text{min}^{-1}, \alpha_2 = 7.69 \times 10^{-2} \text{min}^{-1}) \) where mIU stands for milli-International-Unit, see Veldhuis et al. (1986)). The measurement error has a constant coefficient of variation, namely \( \text{Var}[v_k] = s^2\hat{y}_k^2 \), where \( s = 5.2\% \) is the coefficient of variation. Since spontaneous secretion is studied, \( f(t) \neq 0, t < 0 \), which has been kept into account by letting \( t_0 = -55 \text{ min} \). The regularization parameter was adjusted so as to obtain the same degree of smoothness as in a previous study (De Nicolao et al. 1997) which was achieved for \( \gamma = 449.44 \).

The results are reported in Figure 2. In the left panel, the measured concentrations are plotted together with the continuous-time concentration profile reconstructed by deconvoluting the estimated ISR, which is reported in the right panel. The pulsatile nature of the ISR is apparent. In particular it is easy to distinguish 4-5 major secretory episodes as well as some minor ones. A comparison with Figure 1 in De Nicolao et al. (1997) does not show any appreciable difference with respect to the estimates obtained by discretization. Such a discretization had been carried out on a rather frequent virtual grid (1-min. sampling corresponding to an unknown vector with 250 elements) which entails a significant computational burden. On the contrary, the new deconvolution procedure does not use any discretization and is computationally more efficient.

Concerning this specific application of regularization networks, two possible drawbacks deserve to be men-
tioned. In general, individual impulse response models for each patient are not available, and, as done in our example, an average population model taken from the literature is employed. This raises the issue of robustness of the estimator with respect to interindividual variability of the impulse response. Such a problem has been already investigated in previous studies using a Monte Carlo approach, that is randomly perturbing the population model so as to assess the effect of model uncertainty on the computed estimate (De Nicolao et al. 1997), (Sartorio et al. 1997), (De Nicolao et al. 1999).

Since the only difference with respect to the present paper was the use of (less efficient) discretization methods to compute the regularized solution \( \hat{f} \), the conclusion that the solution is reasonably robust in the face of interindividual variability extends to the present context as well. The second drawback has to do with the intrinsic nonnegativity of concentration values. In this specific set of data the regularization method yielded a nonnegative solution, but in general a constrained optimization method will be needed to minimize the cost functional \( (3) \) (De Nicolao et al. 1997). In any case, it is still helpful to have an efficient algorithm to compute the unconstrained solution that can be used as initialization for an iterative constrained algorithm such as the constrained conjugate gradient.

7 Conclusions

The solution of inverse problems through the regularization method yields a regularization network. This means that the estimate is obtained as a linear combination of activation functions that depend on the linear operator, the regularizing stabilizer, and the sampling grid, but not on the observations.

Once the activation functions are known, the inverse problem can then be solved without any form of numerical integration. Unfortunately, in the general case it is difficult to work out the activation functions in closed form. In the paper, the reconstruction of scalar functions of one variable (time-dependent signals, for instance) from the inversion of finite-dimensional operators is considered. By exploiting the state-space representation and the Bayesian interpretation of regularization it is shown how to compute (without numerical integration) the activation functions of the regularization network as well as the matrices needed to compute the weights. In some specific cases of practical importance, a closed-form expression is provided. Moreover, it is recalled that the weights of the network can be computed in \( O(N) \) operations through Kalman filtering techniques. Finally, the algorithms have been successfully applied to a real-world problem arising in the study of glandular hormone secretion.

Fig. 2. Deconvolution of LH data. Upper panel: measured LH concentrations in plasma (crosses) and reconstructed continuous-time concentration profile (-). Lower panel: estimated Instantaneous Secretion Rate (ISR).

Acknowledgments

This paper has been partially supported by MURST Project “Identification and Control of Industrial Systems”, and ICE-CNR.

References


