1

Continuous-Time Minimum-Mean-Square-Error Filtering

1.1 Introduction

Optimal filtering is concerned with designing the best linear system for recovering data from noisy measurements. It is a model-based approach requiring knowledge of the signal generating system. The signal models, together with the noise statistics are factored into the design in such a way to satisfy an optimality criterion, namely, minimising the square of the error.

A prerequisite technique, the method of least-squares, has its origin in curve fitting. Amid some controversy, Kepler claimed in 1609 that the planets move around the Sun in elliptical orbits [1]. Carl Freidrich Gauss arrived at a better performing method for fitting curves to astronomical observations and predicting planetary trajectories in 1799 [1]. He formally published a least-squares approximation method in 1809 [2], which was developed independently by Adrien-Marie Legendre in 1806 [1]. This technique was famously used by Giusseppe Piazzi to discover and track the asteroid Ceres using a least-squares analysis which was easier than solving Kepler's complicated nonlinear equations of planetary motion [1]. Andrey N. Kolmogorov refined Gauss's theory of least-squares and applied it for the prediction of discrete-time stationary stochastic processes in 1939 [3]. Norbert Wiener, a faculty member at MIT, independently solved analogous continuous-time estimation problems. He worked on defence applications during the Second World War and produced a report entitled *Extrapolation, Interpolation and Smoothing of Stationary Time Series* in 1943. The report was later published as a book in 1949 [4].

Wiener derived two important results, namely, the optimum (non-causal) minimum-mean-square-error solution and the optimum causal minimum-mean-square-error solution [4] – [6]. The optimum causal solution has since become known at the Wiener filter and in the time-invariant case is equivalent to the Kalman filter that was developed subsequently. Wiener pursued practical outcomes and attributed the term "unrealisable filter" to the optimal non-causal solution because "it is not in fact realisable with a finite network of resistances, capacities, and inductances" [4]. Wiener's unrealisable filter is actually the optimum linear smoother.

The optimal Wiener filter is calculated in the frequency domain. Consequently, Section 1.2 touches on some frequency-domain concepts. In particular, the notions of spaces, state-space systems, transfer functions, canonical realisations, stability, causal systems, power spectral density and spectral factorisation are introduced. The Wiener filter is then derived by minimising the square of the error. Three cases are discussed in Section 1.3. First, the

---

"All men by nature desire to know." *Aristotle*

solution to general estimation problem is stated. Second, the general estimation results are specialised to output estimation. The optimal input estimation or equalisation solution is then described. An example, demonstrating the recovery of a desired signal from noisy measurements, completes the chapter.

1.2 Prerequisites

1.2.1 Signals

Consider two continuous-time, real-valued stochastic (or random) signals $v^T(t) = [v_1^T(t),\ v_2^T(t),\ \ldots,\ v_n^T(t)]$, $w^T(t) = [w_1^T(t),\ w_2^T(t),\ \ldots,\ w_n^T(t)]$, with $v_i(t)$, $w_i(t) \in \mathbb{R}$, $i = 1, \ldots n$, which are said to belong to the space $\mathbb{R}^n$, or more concisely $v(t), w(t) \in \mathbb{R}^n$. Let w denote the set of $w(t)$ over all time t, that is, $w = \{ w(t),\ t \in (-\infty, \infty) \}$.

1.2.2 Elementary Functions Defined on Signals

The inner product $\langle v, w \rangle$ of two continuous-time signals v and w is defined by

$$\langle v, w \rangle = \int_{-\infty}^{\infty} v^T w \, dt \ . \tag{1}$$

The 2-norm or Euclidean norm of a continuous-time signal w, $\|w\|_2$, is defined as $\|w\|_2 = \sqrt{\langle w, w \rangle} = \sqrt{\int_{-\infty}^{\infty} w^T w dt}$. The square of the 2-norm, that is, $\|w\|_2^2 = \langle w^T w \rangle = \int_{-\infty}^{\infty} w^T w \, dt$ is commonly known as energy of the signal w.

1.2.3 Spaces

The Lebesgue 2-space, defined as the set of continuous-time signals having finite 2-norm, is denoted by $\mathcal{L}_2$. Thus, $w \in \mathcal{L}_2$ means that the energy of w is bounded. The following properties hold for 2-norms.

(i) $\|v\|_2 = 0 \Rightarrow v = 0$.

(ii) $\|\alpha v\|_2 = |\alpha| \ \|v\|_2$.

(iii) $\|v + w\|_2 \le \|v\|_2 + \|w\|_2$, which is known as the triangle inequality.

(iv) $\|vw\|_2 \le \|v\|_2 \|w\|_2$.

(v) $|\langle v, w \rangle| \le \|v\|_2 \|w\|_2$, which is known as the Cauchy-Schwarz inequality.

See [8] for more detailed discussions of spaces and norms.

"Scientific discovery consists in the interpretation for our own convenience of a system of existence which has been made with no eye to our convenience at all." *Norbert Wiener*

1.2.4 Linear Systems

A linear system is defined as having an output vector which is equal to the value of a linear operator applied to an input vector. That is, the relationships between the output and input vectors are described by linear equations, which may be algebraic, differential or integral. Linear time-domain systems are denoted by upper-case script fonts. Consider two linear systems $\mathcal{G}, \mathcal{H} : \mathbb{R}^p \rightarrow \mathbb{R}^q$, that is, they operate on an input $w \in \mathbb{R}^p$ and produce outputs $\mathcal{G}w, \mathcal{H}w \in \mathbb{R}^q$. The following properties hold.

$$(\mathcal{G} + \mathcal{H})\, w = \mathcal{G}w + \mathcal{H}w, \tag{2}$$

$$(\mathcal{G}\mathcal{H})\, w = \mathcal{G}\,(\mathcal{H}w), \tag{3}$$

$$(\alpha\mathcal{G})\, w = \alpha\,(\mathcal{G}w), \tag{4}$$

where $\alpha \in \mathbb{R}$. An interpretation of (2) is that a parallel combination of $\mathcal{G}$ and $\mathcal{H}$ is equivalent to the system $\mathcal{G} + \mathcal{H}$. From (3), a series combination of $\mathcal{G}$ and $\mathcal{H}$ is equivalent to the system $\mathcal{G}\mathcal{H}$. Equation (4) states that scalar amplification of a system is equivalent to scalar amplification of a system's output.

1.2.5 Polynomial Fraction Systems

The Wiener filtering results [4] – [6] were originally developed for polynomial fraction descriptions of systems which are described below. Consider an n^{th}-order linear, time-invariant system $\mathcal{G}$ that operates on an input $w(t) \in \mathbb{R}$ and produces an output $y(t) \in \mathbb{R}$, that is, $\mathcal{G} : \mathcal{H} : \mathbb{R} \rightarrow \mathbb{R}$. Suppose that the differential equation model for this system is

$$\begin{aligned} a_n \frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \ldots + a_1 \frac{dy(t)}{dt} + a_0 y(t) \\ = b_m \frac{d^m w(t)}{dt^m} + b_{m-1} \frac{d^{m-1} w(t)}{dt^{n-1}} + \ldots + b_1 \frac{dw(t)}{dt} + b_0 w(t), \end{aligned} \tag{5}$$

where $a_0, \ldots a_n$ and $b_0, \ldots b_m$ are real-valued constant coefficients, $a_n \neq 0$, with zero initial conditions. This differential equation can be written in the more compact form

$$\begin{aligned} \left(a_n \frac{d^n}{dt^n} + a_{n-1} \frac{d^{n-1}}{dt^{n-1}} + \ldots + a_1 \frac{d}{dt} + a_0 \right) y(t) \\ = \left(b_m \frac{d^m}{dt^m} + b_{m-1} \frac{d^{m-1}}{dt^{n-1}} + \ldots + b_1 \frac{d}{dt} + b_0 \right) w(t). \end{aligned} \tag{6}$$

1.2.6 The Laplace Transform of a Signal

The two-sided Laplace transform of a continuous-time signal $y(t) \in \mathbb{R}$ is denoted by $Y(s)$ and defined by

$$Y(s) = \int_{-\infty}^{\infty} y(t) e^{-st} dt, \tag{7}$$

"Science is a way of thinking much more than it is a body of knowledge." *Carl Edward Sagan*

where $s = \sigma + j\omega$ is the Laplace transform variable, in which $\sigma, \omega \in \mathbb{R}$ and $j = \sqrt{-1}$. Given a signal $y(t)$ with Laplace transform $Y(s)$, $y(t)$ can be calculated from $Y(s)$ by taking the inverse Laplace Transform of $Y(s)$, which is defined by

$$y(t) = \int_{\sigma - j\infty}^{\sigma + j\infty} Y(s) e^{st} ds \,. \tag{8}$$

Theorem 1 Parseval's Theorem [7]:

$$\int_{-\infty}^{\infty} \left| y(t) \right|^2 dt = \int_{-j\infty}^{j\infty} \left| Y(s) \right|^2 ds \,. \tag{9}$$

Proof. *Let* $y^H(t) = \int_{\sigma - j\infty}^{\sigma + j\infty} Y^H(s) e^{-st} ds$ *and* $Y^H(s)$ *denote the Hermitian transpose (or adjoint) of* $y(t)$ *and* $Y(s)$*, respectively. The left-hand-side of (9) may be written as*

$$\begin{aligned}
\int_{-\infty}^{\infty} \left| y(t) \right|^2 dt &= \int_{-\infty}^{\infty} y^H(t) y(t) dt \\
&= \int_{-j\infty}^{j\infty} \frac{1}{2\pi j} \int_{-\infty}^{\infty} Y^H(s) e^{-st} ds \; y(t) dt \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} y(t) e^{-st} dt \; Y^H(s) ds \\
&= \int_{-j\infty}^{j\infty} Y(s) Y^H(s) ds \\
&= \int_{-j\infty}^{j\infty} \left| Y(s) \right|^2 ds \,.
\end{aligned}$$

□

The above theorem is attributed to Parseval whose original work **[**7] concerned the sums of trigonometric series. An interpretation of (9) is that the energy in the time domain equals the energy in the frequency domain.

1.2.7 Polynomial Fraction Transfer Functions

The steady-state response $y(t) = Y(s)e^{st}$ can be found by applying the complex-exponential input $w(t) = W(s)e^{st}$ to the terms of (6), which results in

$$\left(a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0 \right) Y(s) e^{st} = \left(b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0 \right) W(s) e^{st} \,. \tag{10}$$

Therefore,

$$\begin{aligned}
Y(s) &= \left[\frac{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0} \right] W(s) \\
&= G(s) W(s) \,,
\end{aligned} \tag{11}$$

"No, no, you're not thinking; you're just being logical." *Niels Henrik David Bohr*

where

$$G(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0}. \tag{12}$$

is known as the transfer function of the system. It can be seen from (6) and (12) that the polynomial transfer function coefficients correspond to the system's differential equation coefficients. Thus, knowledge of a system's differential equation is sufficient to identify its transfer function.

1.2.8 Poles and Zeros

The numerator and denominator polynomials of (12) can be factored into m and n linear factors, respectively, to give

$$G(s) = \frac{b_m (s-\beta_1)(s-\beta_2)\ldots(s-\beta_m)}{a_n (s-\alpha_1)(s-\alpha_2)\ldots(s-\alpha_n)}. \tag{13}$$

The numerator of $G(s)$ is zero when $s = \beta_i$, $i = 1 \ldots m$. These values of s are called the zeros of $G(s)$. Zeros in the left-hand-plane are called minimum-phase whereas zeros in the right-hand-plane are called non-minimum phase. The denominator of $G(s)$ is zero when $s = \alpha_i$, $i = 1 \ldots n$. These values of s are called the poles of $G(s)$.

Example 1. Consider a system described by the differential equation $\dot{y}(t) = -y(t) + w(t)$, in which $y(t)$ is the output arising from the input $w(t)$. From (6) and (12), it follows that the corresponding transfer function is given by $G(s) = (s + 1)^{-1}$, which possesses a pole at $s = -1$.

The system in Example 1 operates on a single input and produces a single output, which is known as single-input-single-output (SISO) system. Systems operating on multiple inputs and producing multiple outputs, for example, $\mathcal{G} : \mathbb{R}^p \rightarrow \mathbb{R}^q$, are known as multiple-input-multiple-output (MIMO). The corresponding transfer function matrices can be written as equation (14), where the components $G_{ij}(s)$ have the polynomial transfer function form within (12) or (13).

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) & \ldots & G_{1p}(s) \\ G_{21}(s) & G_{22}(s) & & \\ \vdots & & \ddots & \vdots \\ G_{q1}(s) & & \ldots & G_{qp}(s) \end{bmatrix}. \tag{14}$$

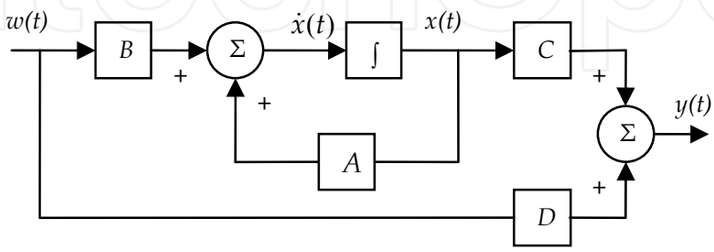


Figure 1. Continuous-time state-space system.

"Nature laughs at the difficulties of integration." *Pierre-Simon Laplace*

1.2.9 State-Space Systems

A system $\mathcal{G}:\mathbb{R}^p \rightarrow \mathbb{R}^q$ having a state-space realisation is written in the form

$$\dot{x}(t) = Ax(t) + Bw(t)\,, \tag{15}$$

$$y(t) = Cx(t) + Dw(t)\,, \tag{16}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{p\times m}$, $C \in \mathbb{R}^{q\times n}$ and $D \in \mathbb{R}^{q\times q}$, in which $w \in \mathbb{R}^p$ is an input, $x \in \mathbb{R}^n$ is a state vector and $y \in \mathbb{R}^q$ is an output. A is known as the state matrix and D is known as the direct feed-through matrix. The matrices B and C are known as the input mapping and the output mapping, respectively. This system is depicted in Fig. 1.

1.2.10 Euler's Method for Numerical Integration

Differential equations of the form (15) could be implemented directly by analog circuits. Digital or software implementations require a method for numerical integration. A first-order numerical integration technique, known as Euler's method, is now derived. Suppose that $x(t)$ is infinitely differentiable and consider its Taylor series expansion in the neighbourhood of t_0

$$\begin{aligned} x(t) &= x(t_0) + \frac{(t-t_0)}{1!}\frac{dx(t_0)}{dt} + \frac{(t-t_0)^2}{2!}\frac{d^2x(t_0)}{dt^2} + \frac{(t-t_0)^3}{3!}\frac{d^3x(t_0)}{dt^3} + \cdots \\ &= x(t_0) + \frac{(t-t_0)}{1!}\dot{x}(t_0) + \frac{(t-t_0)^2}{2!}\ddot{x}(t_0) + \frac{(t-t_0)^3}{3!}\dddot{x}(t_0) + \cdots \end{aligned} \tag{17}$$

Truncating the series after the first order term yields the approximation $x(t) \approx x(t_0) + (t-t_0)\dot{x}(t_0)$. Defining $t_k = t_{k\text{-}1} + \delta_t$ leads to

$$\begin{aligned} x(t_1) &= x(t_0) + \delta_t \dot{x}(t_0) \\ x(t_2) &= x(t_1) + \delta_t \dot{x}(t_1) \\ &\vdots \\ x(t_{k+1}) &= x(t_k) + \delta_t \dot{x}(t_k)\,. \end{aligned} \tag{18}$$

Thus, the continuous-time linear system (15) could be approximated in discrete-time by iterating

$$\dot{x}(t_{k+1}) = Ax(t_k) + Bw(t_k) \tag{19}$$

and (18) provided that δ_t is chosen to be suitably small. Applications of (18) – (19) appear in [9] and in the following example.

"It is important that students bring a certain ragamuffin, barefoot irreverence to their studies; they are not here to worship what is known, but to question it." *Jacob Bronowski*

Example 2. In respect of the continuous-time state evolution (15), consider $A = -1$, $B = 1$ together with the deterministic input $w(t) = \sin(t) + \cos(t)$. The states can be calculated from the known $w(t)$ using (19) and the difference equation (18). In this case, the state error is given by $e(t_k) = \sin(t_k) - x(t_k)$. In particular, root-mean-square-errors of 0.34, 0.031, 0.0025 and 0.00024, were observed for $\delta_t = 1$, 0.1, 0.01 and 0.001, respectively. This demonstrates that the first order approximation (18) can be reasonable when δ_t is sufficiently small.

1.2.11 State-Space Transfer Function Matrix

The transfer function matrix of the state-space system (15) - (16) is defined by

$$G(s) = C(sI - A)^{-1}B + D\,, \tag{20}$$

in which s again denotes the Laplace transform variable.

Example 3. For a state-space model with $A = -1$, $B = C = 1$ and $D = 0$, the transfer function is $G(s) = (s + 1)^{-1}$.

Example 4. For state-space parameters $A = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 5 \end{bmatrix}$ and $D = 0$, the use of Cramer's rule, that is, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, yields the transfer function $G(s) = \frac{(2s + 5)}{(s + 1)(s + 2)} = \frac{1}{(s + 1)} + \frac{1}{(s + 2)}$.

Example 5. Substituting $A = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}$ and $B = C = D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ into (20) results in the transfer function matrix $G(s) = \begin{bmatrix} \frac{s+2}{s+1} & 0 \\ 0 & \frac{s+3}{s+2} \end{bmatrix}$.

1.2.12 Canonical Realisations

The mapping of a polynomial fraction transfer function (12) to a state-space representation (20) is not unique. Two standard state-space realisations of polynomial fraction transfer functions are described below. Assume that: the transfer function has been expanded into the sum of a direct feed-though term plus a strictly proper transfer function, in which the order of the numerator polynomial is less than the order of the denominator polynomial; and the strictly proper transfer function has been normalised so that $a_n = 1$. Under these assumptions, the system can be realised in the controllable canonical form which is parameterised by [10]

"Science is everything we understand well enough to explain to a computer. Art is everything else."
David Knuth

$$A=\begin{bmatrix} -a_{n-1} & -a_{n-2} & \dots & -a_1 & -a_0 \\ 1 & 0 & & \dots & 0 \\ 0 & 1 & & & \\ \vdots & & \ddots & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, B=\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \text{ and } C=\begin{bmatrix} b_m & b_{m-1} & \dots & b_1 & b_0 \end{bmatrix}.$$

The system can be also realised in the observable canonical form which is parameterised by

$$A=\begin{bmatrix} -a_{n-1} & 1 & 0 & \dots & 0 \\ -a_{n-2} & 0 & 1 & & 0 \\ \vdots & & & \ddots & 0 \\ -a_1 & & & 0 & 1 \\ -a_0 & 0 & \dots & 0 & 0 \end{bmatrix}, B=\begin{bmatrix} b_m \\ b_{m-1} \\ \vdots \\ b_1 \\ b_0 \end{bmatrix} \text{ and } C=\begin{bmatrix} 1 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

1.2.13 Asymptotic Stability

Consider a continuous-time, linear, time-invariant n^{th}-order system $\mathcal{G}$ that operates on an input w and produces an output y. The system $\mathcal{G}$ is said to be asymptotically stable if the output remains bounded, that is, $y \in \mathcal{L}_2$, for any $w \in \mathcal{L}_2$. This is also known as bounded-input-bounded-output stability. Two equivalent conditions for $\mathcal{G}$ to be asymptotically stable are:

- The real part of the eigenvalues of the system's state matrix are in the left-hand-plane, that is, for A of (20), $\text{Re}\{\lambda_i(A)\} < 0$, $i = 1 \dots n$.
- The real part of the poles of the system's transfer function are in the left-hand-plane, that is, for α_i of (13), $\text{Re}\{\alpha_i\} < 0$, $i = 1 \dots n$.

Example 6. A state-space system having $A = -1$, $B = C = 1$ and $D = 0$ is stable, since $\lambda(A) = -1$ is in the left-hand-plane. Equivalently, the corresponding transfer function $G(s) = (s + 1)^{-1}$ has a pole at $s = -1$ which is in the left-hand-plane and so the system is stable. Conversely, the transfer function $G^T(-s) = (1 - s)^{-1}$ is unstable because it has a singularity at the pole $s = 1$ which is in the right hand side of the complex plane. $G^T(-s)$ is known as the adjoint of $G(s)$ which is discussed below.

1.2.14 Adjoint Systems

An important concept in the ensuing development of filters and smoothers is the adjoint of a system. Let $\mathcal{G}: \mathbb{R}^p \rightarrow \mathbb{R}^q$ be a linear system operating on the interval [0, T]. Then $\mathcal{G}^H: \mathbb{R}^q \rightarrow \mathbb{R}^p$, the adjoint of $\mathcal{G}$, is the unique linear system such that $\langle y, \mathcal{G} w \rangle = \langle \mathcal{G}^H y, w \rangle$, for all $y \in \mathbb{R}^q$ and $w \in \mathbb{R}^p$. The following derivation is a simplification of the time-varying version that appears in [11].

"Science might almost be redefined as the process of substituting unimportant questions which can be answered for important questions which cannot." *Kenneth Ewart Boulding*

Lemma 1 (State-space representation of an adjoint system): *Suppose that a continuous-time linear time-invariant system* $\mathcal{G}$ *is described by*

$$\dot{x}(t) = Ax(t) + Bw(t)\ , \tag{21}$$

$$y(t) = Cx(t) + Dw(t)\ , \tag{22}$$

with $x(t_0) = 0$. *The adjoint* $\mathcal{G}^H$ *is the linear system having the realisation*

$$\dot{\zeta}(t) = -A^T\zeta(t) - C^T u(t)\ , \tag{23}$$

$$z(t) = B^T\zeta(t) + D^T u(t)\ , \tag{24}$$

with $\zeta(T) = 0$.

Proof: *The system (21) – (22) can be written equivalently*

$$\begin{bmatrix} \frac{d}{dt}I - A & -B \\ C & D \end{bmatrix}\begin{bmatrix} x(t) \\ w(t) \end{bmatrix} = \begin{bmatrix} 0(t) \\ y(t) \end{bmatrix} \tag{25}$$

with $x(t_0) = 0$. *Thus*

$$\begin{aligned} <y,\ \mathcal{G}\, w> &= \left\langle \begin{bmatrix} \zeta \\ u \end{bmatrix}, \begin{bmatrix} \frac{d}{dt}I - A & -B \\ C & D \end{bmatrix}\begin{bmatrix} x \\ w \end{bmatrix} \right\rangle \\ &= \int_o^T \left(\zeta^T \frac{dx}{dt} \right) dt - \int_0^T \zeta^T (Ax + Bw)\ dt + \int_0^T u^T (Cx + Dw)\ dt\ . \end{aligned} \tag{26}$$

Integrating the last term by parts gives

$$\begin{aligned} <y,\ \mathcal{G}\, w> &= \zeta^T(T)x(T) - \int_0^T \left(\frac{d\zeta^T}{dt} \right) x\ dt - \int_0^T \zeta^T (Ax + Bw)\ dt\ . \\ &\quad + \int_0^T u^T (Cx + Dw)\ dt \\ &= \left\langle \begin{bmatrix} -\left(\frac{d}{dt}I - A^T \right) & C^T \\ -B^T & D^T \end{bmatrix}\begin{bmatrix} \zeta \\ u \end{bmatrix}, \begin{bmatrix} x \\ w \end{bmatrix} \right\rangle + \lambda^T(T)x(T) \\ &= <\mathcal{G}^H \boldsymbol{y}, \boldsymbol{w}>, \end{aligned} \tag{27}$$

where $\mathcal{G}^H$ *is given by (23) – (24).* □

"If you thought that science was certain—well, that is just an error on your part." *Richard Phillips Feynman*

Thus, the adjoint of a system having the parameters $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is a system with $\begin{bmatrix} -A^T & -C^T \\ B^T & D^T \end{bmatrix}$.

Adjoint systems have the property $(\mathcal{G}^H)^H = \mathcal{G}$. The adjoint of the transfer function matrix $G(s)$ is denoted as $G^H(s)$ and is defined by the transfer function matrix

$$G^H(s) = G^T(-s). \tag{28}$$

Example 7. Suppose that a system $\mathcal{G}$ has state-space parameters $A = -1$ and $B = C = D = 1$. From (23) – (24), an adjoint system has the state-space parameters $A = 1$, $B = D = 1$ and $C = -1$ and the corresponding transfer function is $G^H(s) = 1 - (s - 1)^{-1} = (- s + 2)(- s + 1)^{-1} = (s - 2)(s - 1)^{-1}$, which is unstable and non-minimum-phase. Alternatively, the adjoint of $G(s) = 1 + (s + 1)^{-1} = (s + 2)(s + 1)^{-1}$ can be obtained using (28), namely $G^H(s) = G^T(-s) = (- s + 2)(- s + 1)^{-1}$.

1.2.15 Causal and Noncausal Systems

A causal system is a system that depends exclusively on past and current inputs.

Example 8. The differential of $x(t)$ with respect to t is defined by $\dot{x}(t) = \lim_{dt \to 0} \frac{x(t + dt) - x(t)}{dt}$.

Consider

$$\dot{x}(t) = Ax(t) + Bw(t) \tag{29}$$

with $\mathrm{Re}\{\lambda_i(A)\} < 0$, $i = 1, \ldots, n$. The positive sign of $\dot{x}(t)$ within (29) denotes a system that proceeds forward in time. This is called a causal system because it depends only on past and current inputs.

Example 9. The negative differential of $\xi(t)$ with respect to t is defined by $-\dot{\xi}(t) = \lim_{dt \to 0} \frac{\xi(t) - \xi(t + dt)}{dt}$. Consider

$$-\dot{\zeta}(t) = A^T\zeta(t) + C^Tu(t) \tag{30}$$

with $\mathrm{Re}\{\lambda_i(A)\} = \mathrm{Re}\{\lambda_i(A^T)\} < 0$, $i = 1 \ldots n$. The negative sign of $\dot{\zeta}(t)$ within (30) denotes a system that proceeds backwards in time. Since this system depends on future inputs, it is termed noncausal. Note that $\mathrm{Re}\{\lambda_i(A)\} < 0$ implies $\mathrm{Re}\{\lambda_i(-A)\} > 0$. Hence, if causal system (21) – (22) is stable, then its adjoint (23) – (24) is unstable.

1.2.16 Realising Unstable System Components

Unstable systems are termed unrealisable because their outputs are not in $\mathcal{L}_2$ that is, they are unbounded. In other words, they cannot be implemented as forward-going systems. It follows from the above discussion that an unstable system component can be realised as a stable noncausal or backwards system.

Suppose that the time domain system $\mathcal{G}$ is stable. The adjoint system $z = \mathcal{G}^H u$ can be realised by the following three-step procedure.

"We haven't the money, so we've got to think." *Baron Ernest Rutherford*

- Time-reverse the input signal $u(t)$, that is, construct $u(\tau)$, where $\tau = T - t$ is a time-to-go variable (see [12]).
- Realise the stable system $\mathcal{G}^T$

$$\dot{\zeta}(\tau) = A^T\zeta(\tau) + C^T u(\tau)\,, \tag{31}$$

$$z(\tau) = B^T\zeta(\tau) + D^T u(\tau)\,, \tag{32}$$

with $\zeta(T) = 0$.
- Time-reverse the output signal $z(\tau)$, that is, construct $z(t)$.

The above procedure is known as noncausal filtering or smoothing; see the discrete-time case described in [13]. Thus, a combination of causal and non-causal system components can be used to implement an otherwise unrealisable system. This approach will be exploited in the realisation of smoothers within subsequent sections.

Example 10. Suppose that it is required to realise the unstable system $G(s) = G_2^H(s)G_1(s)$ over an interval [0, T], where $G_1(s) = (s+1)^{-1}$ and $G_2(s) = (s+2)^{-1}$. This system can be realised using the processes shown in Fig. 2.

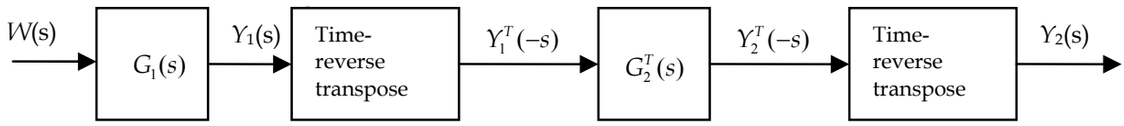


Figure 2. Realising an unstable $G(s) = G_2^H(s)G_1(s)$.

1.2.17 Power Spectral Density

The power of a voltage signal applied to a 1-ohm load is defined as the squared value of the signal and is expressed in watts. The power spectral density is expressed as power per unit bandwidth, that is, W/Hz. Consider again a linear, time-invariant system $y = \mathcal{G}w$ and its corresponding transfer function matrix $G(s)$. Assume that w is a zero-mean, stationary, white noise process with $E\{w(t)w^T(\tau)\} = Q\delta(t-\tau)$, in which δ denotes the Dirac delta function. Then $\Phi_{yy}(s)$, the power spectral density of y, is given by

$$\Phi_{yy}(s) = GQG^H(s)\,, \tag{33}$$

which has the property $\Phi_{yy}(s) = \Phi_{yy}(-s)$.

The total energy of a signal is the integral of the power of the signal over time and is expressed in watt-seconds or joules. From Parseval's theorem (9), the average total energy of $y(t)$ is

$$\int_{-j\infty}^{j\infty}\Phi_{yy}(s)ds = \int_{-\infty}^{\infty}\left|y(t)\right|^2 dt = \left\|y(t)\right\|_2^2 = E\{y^T(t)y(t)\}\,, \tag{34}$$

which is equal to the area under the power spectral density curve.

"Time is what prevents everything from happening at once." *John Archibald Wheeler*

1.2.18 Spectral Factorisation

Suppose that noisy measurements

$$z(t) = y(t) + v(t) \tag{35}$$

of a linear, time-invariant system $\mathcal{G}$, described by (21) - (22), are available, where $v(t) \in \mathbb{R}^q$ is an independent, zero-mean, stationary white noise process with $E\{v(t)v^T(\tau)\} = R\delta(t-\tau)$. Let

$$\Phi_{zz}(s) = GQG^H(s) + R \tag{36}$$

denote the spectral density matrix of the measurements $z(t)$. Spectral factorisation was pioneered by Wiener (see [4] and [5]). It refers to the problem of decomposing a spectral density matrix into a product of a stable, minimum-phase matrix transfer function and its adjoint. In the case of the output power spectral density (36), a spectral factor $\Delta(s)$ satisfies $\Delta(s)\Delta^H(s) = \Phi_{zz}(s)$.

The problem of spectral factorisation within continuous-time Wiener filtering problems is studied in [14]. The roots of the transfer function polynomials need to be sorted into those within the left-hand-plane and the right-hand plane. This is an eigenvalue decomposition problem – see the survey of spectral factorisation methods detailed in [11].

Example 11. In respect of the observation spectral density (36), suppose that $G(s) = (s + 1)^{-1}$ and $Q = R = 1$, which results in $\Phi_{zz}(s) = (- s^2 + 2)(- s^2 + 1)^{-1}$. By inspection, the spectral factor $\Delta(s) = (s+\sqrt{2})(s+1)^{-1}$ is stable, minimum-phase and satisfies $\Delta(s)\Delta^H(s) = \Phi_{zz}(s)$.

1.3 Minimum-Mean-Square-Error Filtering

1.3.1 Filter Derivation

Now that some underlying frequency-domain concepts have been introduced, the Wiener filter [4] – [6] can be described. A Wiener-Hopf derivation of the Wiener filter appears in [4], [6]. This section describes a simpler completing-the-square approach (see [14], [16]). Consider a stable linear time-invariant system having a transfer function matrix $G_2(s) = C_2(sI - A)^{-1}B + D_2$. Let $Y_2(s)$, $W(s)$, $V(s)$ and $Z(s)$ denote the Laplace transforms of the system's output, measurement noise, process noise and observations, respectively, so that

$$Z(s) = Y_2(s) + V(s). \tag{37}$$

Consider also a fictitious reference system having the transfer function $G_1(s) = C_1(sI - A)^{-1}B + D_1$ as shown in Fig. 3. The problem is to design a filter transfer function $H(s)$ to calculate estimates $\hat{Y}_1(s) = H(s)Z(s)$ of $Y_1(s)$ so that the energy $\int_{-j\infty}^{j\infty} E(s)E^H(s)ds$ of the estimation error

$$E(s) = Y_1(s) - \hat{Y}_1(s) \tag{38}$$

is minimised.

"Science may be described as the art of systematic over-simplification." *Karl Raimund Popper*

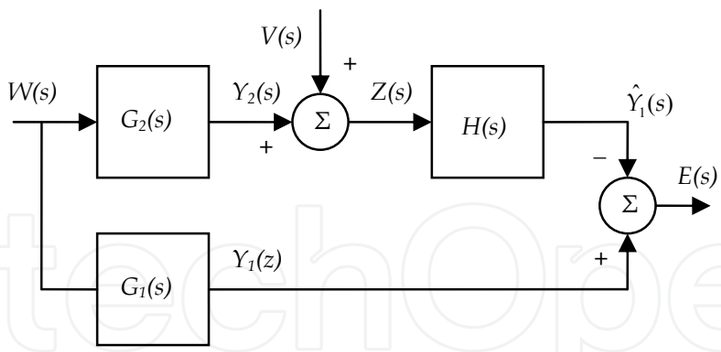


Figure 3. The s-domain general filtering problem.

It follows from Fig. 3 that $E(s)$ is generated by

$$E(s) = -\begin{bmatrix} H(s) & HG_2(s) - G_1(s) \end{bmatrix} \begin{bmatrix} V(s) \\ W(s) \end{bmatrix}. \tag{39}$$

The error power spectrum density matrix is denoted by $\Phi_{ee}(s)$ and given by the covariance of $E(s)$, that is,

$$\begin{aligned} \Phi_{ee}(s) &= E(s)E^H(s) \\ &= \begin{bmatrix} H(s) & HG_2(s) - G_1(s) \end{bmatrix} \begin{bmatrix} R & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} H^H(s) \\ G_2^H H^H(s) - G_1^H(s) \end{bmatrix} \\ &= G_1 Q G_1^H(s) - G_1 Q G_2^H H^H(s) - H G_2 Q G_1^H(s) + H \Delta \Delta^H H^H(z), \end{aligned} \tag{40}$$

where

$$\Delta \Delta^H(s) = G_2 Q G_2^H(s) + R \tag{41}$$

is the spectral density matrix of the measurements. The quantity $\Delta(s)$ is a spectral factor, which is unique up to the product of an inner matrix. Denote $\Delta^{-H}(s) = (\Delta^H)^{-1}(s)$. Completing the square within (40) yields

$$\begin{aligned} \Phi_{ee}(s) &= G_1 Q G_1^H(s) - G_1 Q G_2^H (\Delta \Delta^H)^{-1} G_2 Q G_1^H(s) \\ &\quad + (H\Delta(s) - G_1 Q G_2^H \Delta^{-H}(s))(H\Delta(s) - G_1 Q G_2^H \Delta^{-H}(s))^H. \end{aligned} \tag{42}$$

It follows that the total energy of the error signal is given by

$$\begin{aligned} \int_{-j\infty}^{j\infty} \Phi_{ee}(s) ds &= \int_{-j\infty}^{j\infty} G_1 Q G_1^H(s) - G_1 Q G_2^H (\Delta \Delta^H)^{-1} G_2 Q G_1^H(s) ds \\ &\quad + \int_{-j\infty}^{j\infty} (H\Delta(s) - G_1 Q G_2^H \Delta^{-H}(s))(H\Delta(s) - G_1 Q G_2^H \Delta^{-H}(s))^H ds. \end{aligned} \tag{43}$$

"Science is what you know. Philosophy is what you don't know." *Earl Bertrand Arthur William Russell*

The first term on the right-hand-side of (43) is independent of $H(s)$ and represents a lower bound of $\int_{-j\infty}^{j\infty} \Phi_{ee}(s)ds$. The second term on the right-hand-side of (43) may be minimised by a judicious choice for $H(s)$.

Theorem 2: *The above linear time-invariant filtering problem with by the measurements (37) and estimation error (38) has the solution*

$$H(s) = G_1 Q G_2^H \Delta^{-H} \Delta^{-1}(s) . \tag{44}$$

which minimises $\int_{-j\infty}^{j\infty} \Phi_{ee}(s)ds$.

Proof*: The result follows by setting* $H\Delta(s) - G_1 Q G_2^H \Delta^{-H}(s) = 0$ *within (43).* □

By Parseval's theorem, the minimum mean-square-error solution (44) also minimises $\|e(t)\|_2^2$.

The solution (44) is unstable because the factor $G_2^H(\Delta^H)^{-1}(s)$ possesses right-hand-plane poles. This optimal noncausal solution is actually a smoother, which can be realised by a combination of forward and backward processes. Wiener called (44) the optimal unrealisable solution because it cannot be realised by a memory-less network of capacitors, inductors and resistors [4].

The transfer function matrix of a realisable filter is given by

$$H(s) = \left\{ G_1 Q G_2^H (\Delta^H)^{-1} \right\}_+ \Delta^{-1}(s) , \tag{45}$$

in which $\{\}_+$ denotes the causal part. A procedure for finding the causal part of a transfer function is described below.

1.3.2 Finding the Causal Part of a Transfer Function

The causal part of transfer function can be found by carrying out the following three steps.

- If the transfer function is not strictly proper, that is, if the order of the numerator is not less than the degree of the denominator, then perform synthetic division to isolate the constant term.
- Expand out the (strictly proper) transfer function into the sum of stable and unstable partial fractions.
- The causal part is the sum of the constant term and the stable partial fractions.

Incidentally, the noncausal part is what remains, namely the sum of the unstable partial fractions.

Example 12. Consider $G(s) = (s^2 - \beta^2)(s^2 - \alpha^2)^{-1}$ with $\alpha, \beta < 0$. Since $G_2(s)$ possesses equal order numerator and denominator polynomials, synthetic division is required, which yields $G_2(s) = 1 + (\alpha^2 - \beta^2)(s^2 - \alpha^2)^{-1}$. A partial fraction expansion results in

---

"There is an astonishing imagination, even in the science of mathematics." *Francois-Marie Arouet de Voltaire*

$$\frac{(\alpha^2-\beta^2)}{(s^2-\alpha^2)} = \frac{0.5\alpha^{-1}(\alpha^2-\beta^2)}{(s+\alpha)} - \frac{0.5\alpha^{-1}(\alpha^2-\beta^2)}{(s-\alpha)}.$$

Thus, the causal part of $G(s)$ is $\{G(s)\}_+ = 1 - 0.5\alpha^{-1}(\alpha^2-\beta^2)(s-\alpha)^{-1}$. The noncausal part of $G(s)$ is denoted as $\{G(s)\}_-$ and is given by $\{G(s)\}_- = 0.5\alpha^{-1}(\alpha^2-\beta^2)(s-\alpha)^{-1}$. It is easily verified that $G(s) = \{G(s)\}_+ + \{G(s)\}_-$.

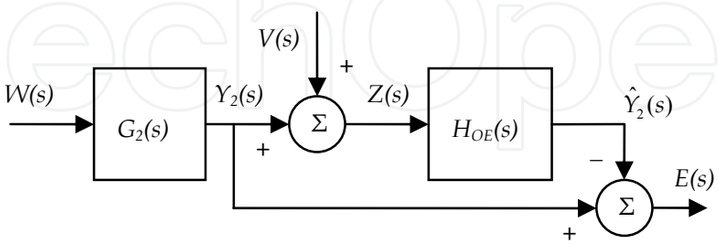


Figure 4. The s-domain output estimation problem.

1.3.3 Minimum-Mean-Square-Error Output Estimation

In output estimation, the reference system is the same as the generating system, as depicted in Fig. 4. The simplification of the optimal noncausal solution (44) of Theorem 2 for the case $G_1(s) = G_2(s)$ can be expressed as

$$\begin{aligned} H_{OE}(s) &= G_2 Q G_2^H \Delta^{-H}\Delta^{-1}(s) \\ &= G_2 Q G_2^H (\Delta\Delta^H)^{-1}(s) \\ &= (\Delta\Delta^H - R)(\Delta\Delta^H)^{-1}(s) \\ &= I - R\Delta^{-H}\Delta^{-1}(s). \end{aligned} \tag{46}$$

The optimal causal solution for output estimation is

$$\begin{aligned} H_{OE}(s) &= \left\{G_2 Q G_2^H \Delta^{-H}\right\}_+ \Delta^{-1}(s) \\ &= I - R\left\{\Delta^{-H}\right\}_+ \Delta^{-1}(s) \\ &= I - R^{1/2}\Delta^{-1}(s). \end{aligned} \tag{47}$$

When the measurement noise becomes negligibly small, the output estimator approaches a short circuit, that is,

$$\lim_{R\to 0, s\to 0} \left|H_{OE}(s)\right| = I. \tag{48}$$

"Science is the topography of ignorance." *Oliver Wendell Holmes*

The observation (48) can be verified by substituting $\Delta\Delta^H(s) = G_2QG_2^H(s)$ into (46). This observation is consistent with intuition, that is, when the measurements are perfect, filtering will be superfluous.

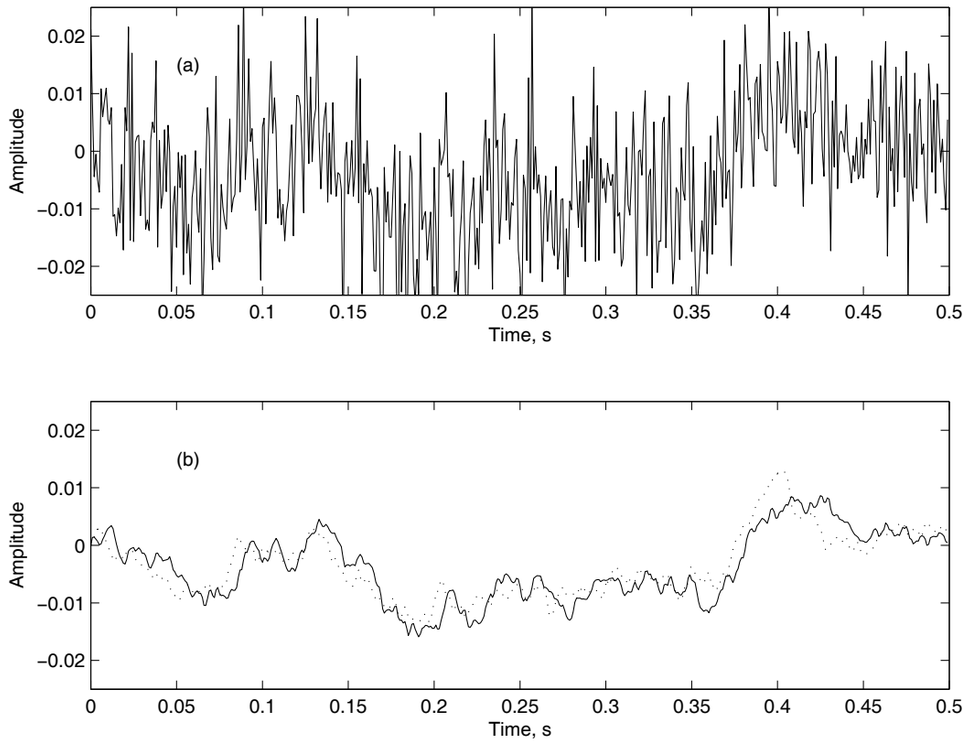


Figure 5. Sample trajectories for Example 13: (a) measurement, (b) system output (dotted line) and filtered signal (solid line).

Example 13. Consider a scalar output estimation problem, where $G_2(s) = (s-\alpha)^{-1}$, $\alpha = -1$, $Q = 1$ and $R = 0.0001$. Then $G_2QG_2^H(s) = Q(-s^2+\alpha^2)^{-1}$ and $\Delta\Delta^H(s) = -(Rs^2 + R\alpha^2 + Q)(-s^2+\alpha^2)^{-1}$, which leads to $\Delta(s) = R^{1/2}(s + \sqrt{\alpha^2 + Q/R})(s-\alpha)^{-1}$. Therefore, $G_2QG_2^H(\Delta^H)^{-1}(s) = \frac{Q}{(s-\alpha)(-s-\alpha)} \frac{(-s-\alpha)}{R^{1/2}(-s+\sqrt{\alpha^2+Q/R})} = \frac{Q}{R^{1/2}(s-\alpha)(-s+\sqrt{\alpha^2+Q/R})}$, in which a common pole and zero were cancelled. Expanding into partial fractions and taking the causal part results in

$$\{G_2QG_2^H(\Delta^H)^{-1}(s)\}_+ = \frac{\left.\frac{Q}{R^{1/2}(-s+\sqrt{\alpha^2+Q/R})}\right|_{s=\alpha}}{(s-\alpha)}$$

"Science is the systematic classification of experience." *George Henry Lewes*

and

$$H_{OE}(s) = \{G_2 Q G_2^H \Delta^{-H}\}_+ \Delta^{-1}(s) = \frac{\alpha + \sqrt{\alpha^2 + Q/R}\,)}{s + \sqrt{\alpha^2 + Q/R}\,)}.$$

Substituting $\alpha = -1$, Q = 1 and R = 0.0001 yields $H(s) = 99(s+100)^{-1}$. By inspection, $\lim_{s \to 0} |H(s)| = \frac{99}{100}$, which illustrates the low measurement noise asymptote (48). Some sample trajectories from a simulation conducted with δ_t = 0.001 s are shown in Fig. 5. The input measurements are shown in Fig. 5(a). It can be seen that the filtered signal (the solid line of Fig. 5 (b)) estimates the system output (the dotted line of Fig. 5(b)).

1.3.4 Minimum-Mean-Square-Error Input Estimation

In input estimation problems, it is desired to estimate the input process $w(t)$, as depicted in Fig. 6. This is commonly known as an equalisation problem, in which it is desired to mitigate the distortion introduced by a communication channel $G_2(s)$. The simplification of the general noncausal solution (44) of Theorem 2 for the case of $G_2(s) = I$ results in

$$H_{IE}(s) = Q G_2^H \Delta^{-H} \Delta^{-1}(s). \tag{49}$$

Equation (49) is known as the optimum minimum-mean-square-error noncausal equaliser [12]. Assume that: $G_2(s)$ is proper, that is, the order of the numerator is the same as the order of the denominator, and the zeros of $G_2(s)$ are in the left-hand-plane. Under these conditions, when the measurement noise becomes negligibly small, the equaliser estimates the inverse of the system model, that is,

$$\lim_{R \to 0} H_{IE}(s) = G_2^{-1}(s). \tag{50}$$

The observation (50) can be verified by substituting $\Delta\Delta^H(s) = G_2 Q G_2^H(s)$ into (49). In other words, if the channel model is invertible and signal to noise ratio is sufficiently high, the equaliser will estimate $w(t)$. When measurement noise is present the equaliser no longer approximates the channel inverse because some filtering is also required. In the limit, when the signal to noise ratio is sufficiently low, the equaliser approaches an open circuit, namely,

$$\lim_{Q \to 0, s \to 0} |H_{IE}(s)| = 0. \tag{51}$$

The observation (51) can be verified by substituting Q = 0 into (49). Thus, when the equalisation problem is dominated by measurement noise, the estimation error is minimised by ignoring the data.

"All of the biggest technological inventions created by man - the airplane, the automobile, the computer - says little about his intelligence, but speaks volumes about his laziness." *Mark Raymond Kennedy*

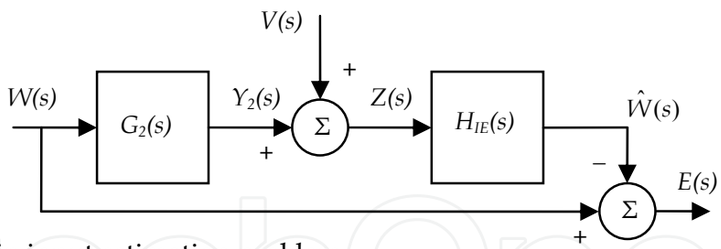


Figure 6. The s-domain input estimation problem.

1.4 Conclusion

Continuous-time, linear, time-invariant systems can be described via either a differential equation model or as a state-space model. Signal models can be written in the time-domain as

$$\left(a_n \frac{d^n}{dt^n} + a_{n-1}\frac{d^{n-1}}{dt^{n-1}} + \ldots + a_1 \frac{d}{dt} + a_0\right) y(t) = \left(b_m \frac{d^m}{dt^m} + b_{m-1}\frac{d^{m-1}}{dt^{m-1}} + \ldots + b_1 \frac{d}{dt} + b_0\right) w(t) .$$

Under the time-invariance assumption, the system transfer function matrices exist, which are written as polynomial fractions in the Laplace transform variable

$$Y(s) = \left[\frac{b_m s^m + b_{m-1}s^{m-1} + \ldots + b_1 s + b_0}{a_n s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0}\right] W(s) = G(s)W(s).$$

Thus, knowledge of a system's differential equation is sufficient to identify its transfer function. If the poles of a system's transfer function are all in the left-hand-plane then the system is asymptotically stable. That is, if the input to the system is bounded then the output of the system will be bounded.

The optimal solution minimises the energy of the error in the time domain. It is found in the frequency domain by minimising the mean-square-error. The main results are summarised in Table 1. The optimal noncausal solution has unstable factors. It can only be realised by a combination of forward and backward processes, which is known as smoothing. The optimal causal solution is also known as the Wiener filter.

In output estimation problems, $C_1 = C_2$, $D_1 = D_2$, that is, $G_1(s) = G_2(s)$ and when the measurement noise becomes negligible, the solution approaches a short circuit. In input estimation or equalisation, $C_1 = 0$, $D_1 = I$, that is, $G_1(s) = I$ and when the measurement noise becomes negligible, the optimal equaliser approaches the channel inverse, provided the inverse exists. Conversely, when the problem is dominated by measurement noise then the equaliser approaches an open circuit.

"Read Euler, read Euler, he is our master in everything." *Pierre-Simon Laplace*

	ASSUMPTIONS	MAIN RESULTS
Signals and systems	$E\{w(t)\} = E\{W(s)\} = E\{\mathrm{v}(t)\} = E\{V(s)\} = 0$. $E\{w(t)w^T(t)\} = E\{W(s)W^T(s)\} = Q > 0$ and $E\{\mathrm{v}(t)v^T(t)\} = E\{V(s)V^T(s)\} = R > 0$ are known. A, B, C_1, C_2, D_1 and D_2 are known. $G_1(s)$ and $G_2(s)$ are stable, *i.e.*, $Re\{\lambda_i(A)\} < 0$.	$G_1(s) = C_1(sI - A)^{-1}B + D_1$ $G_2(s) = C_2(sI - A)^{-1}B + D_2$
Spectral factorisation	$\Delta(s)$ and $\Delta^{-1}(s)$ are stable, *i.e.*, the poles and zeros of $\Delta(s)$ are in the left-half-plane.	$\Delta\Delta^H(s) = G_2QG_2^H(s) + R$
Non-causal solution		$H(s) = G_1QG_2^H(\Delta^H)^{-1}\Delta^{-1}(s)$
Causal solution		$H(s) = \left\{G_1QG_2^H(\Delta^H)^{-1}\right\}_+ \Delta^{-1}(s)$

Table 1. Main results for the continuous-time general filtering problem.

1.5 Problems

Problem 1. Find the transfer functions and comment on stability of the systems having the following polynomial fractions.

(a) $\dddot{y} + 7\ddot{y} + 12\dot{y} = \ddot{w} + \dot{w} - 2w$.

(b) $\dddot{y} + 1\ddot{y} - 20\dot{y} = \ddot{w} + 5\dot{w} + 6w$.

(c) $\dddot{y} + 11\ddot{y} + 30\dot{y} = \ddot{w} - 7\dot{w} + 12w$.

(d) $\dddot{y} - 13\ddot{y} + 42\dot{y} = \ddot{w} + 9\dot{w} + 20w$.

(e) $\dddot{y} - 15\ddot{y} + 56\dot{y} = \ddot{w} + 11\dot{w} + 30w$.

Problem 2. Find the transfer functions and comment on the stability for systems having the following state-space parameters.

(a) $A = \begin{bmatrix} -7 & -12 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $C = \begin{bmatrix} -6 & -14 \end{bmatrix}$ and $D = 1$.

"The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them." *William Henry Bragg*

(b) $A=\begin{bmatrix}-7 & 20\\ 1 & 0\end{bmatrix}$, $B=\begin{bmatrix}1\\ 0\end{bmatrix}$, $C=\begin{bmatrix}-2 & 26\end{bmatrix}$ and $D=1$.

(c) $A=\begin{bmatrix}-11 & -30\\ 1 & 0\end{bmatrix}$, $B=\begin{bmatrix}1\\ 0\end{bmatrix}$, $C=\begin{bmatrix}-18 & -18\end{bmatrix}$ and $D=1$.

(d) $A=\begin{bmatrix}13 & -42\\ 1 & 0\end{bmatrix}$, $B=\begin{bmatrix}1\\ 0\end{bmatrix}$, $C=\begin{bmatrix}22 & -22\end{bmatrix}$ and $D=1$.

(e) $A=\begin{bmatrix}-15 & -56\\ 1 & 0\end{bmatrix}$, $B=\begin{bmatrix}1\\ 0\end{bmatrix}$, $C=\begin{bmatrix}-4 & -26\end{bmatrix}$ and $D=1$.

Problem 3. Calculate the spectral factors for $\Phi_{zz}(s)=GQG^H(s)+R$ having the following models and noise statistics.

(a) $G(s)=(s+1)^{-1}$, Q = 2 and R = 1.

(b) $G(s)=(s+2)^{-1}$, Q = 5 and R = 1.

(c) $G(s)=(s+3)^{-1}$, Q = 7 and R = 1.

(d) $G(s)=(s+4)^{-1}$, Q = 9 and R = 1.

(e) $G(s)=(s+5)^{-1}$, Q = 11 and R = 1.

Problem 4. Calculate the optimal causal output estimators for Problem 3.

Problem 5. Consider the error spectral density matrix

$$\Phi_{ee}(s)=[H\Delta-G_1QG_2^H(\Delta^H)^{-1}][H\Delta-G_1QG_2^H(\Delta^H)^{-1}]^H(s)$$

$$+[G_1QG_1^H-G_1QG_2^H(\Delta\Delta^H)^{-1}G_2QG_1^H](s)\,.$$

(a) Derive the optimal output estimator.

(b) Derive the optimal causal output estimator.

(c) Derive the optimal input estimator.

"Nothing shocks me. I'm a scientist." *Harrison Ford*

Problem 6 [16**].** In respect of the configuration in Fig. 2, suppose that $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 25 \\ 25 \\ -25 \end{bmatrix}$, $C_2 = \begin{bmatrix} -1 & 2 & 1 \end{bmatrix}$, $C_1 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$, $D = 0$, $Q = 1$ and $R = 1$. Show that the optimal causal filter is given by $H(s) = -(16.9s^2 + 86.5s + 97.3)(s^3 + 8.64s^2 + 30.3s + 50.3)^{-1}$.

Problem 7 [18**].** Suppose that $G_2 Q G_2^H(s) = \dfrac{-3600}{s^2(169 - s^2)}$ and $R(s) = 1$. Show that the optimal causal filter for output estimation is given by $H_{OE}(s) = (4s + 60)(s^2 + 17s + 60)^{-1}$.

1.6 Glossary

The following terms have been introduced within this section.

$\mathbb{R}$	The space of real numbers.
$\mathbb{R}^n$	The space of real-valued n-element column vectors.
t	The real-valued continuous-time variable. For example, $t \in (-\infty, \infty)$ and $t \in [0, \infty)$ denote $-\infty < t < \infty$ and $0 \le t < \infty$, respectively.
$w(t) \in \mathbb{R}^n$	A continuous-time, real-valued, n-element stationary stochastic input signal.
w	The set of $w(t)$ over a prescribed interval.
$\mathcal{G} : \mathbb{R}^p \longrightarrow \mathbb{R}^q$	A linear system that operates on a p-element input signal and produces a q-element output signal.
$y = \mathcal{G}w$	The output of a linear system $\mathcal{G}$ that operates on an input signal w.
A, B, C, D	Time-invariant state space matrices of appropriate dimension. The system $\mathcal{G}$ is assumed to have the realisation $\dot{x}(t) = Ax(t) + Bw(t)$, $y(t) = Cx(t) + Dw(t)$ in which $w(t)$ is known as the process noise or input signal.
$v(t)$	A stationary stochastic measurement noise signal.
$\delta(t)$	The Dirac delta function.
Q and R	Time-invariant covariance matrices of stochastic signals $w(t)$ and $v(t)$, respectively.
s	The Laplace transform variable.
$Y(s)$	The Laplace transform of a continuous-time signal $y(t)$.
$G(s)$	The transfer function matrix of a system $\mathcal{G}$. For example, the transfer function matrix of the system $\dot{x}(t) = Ax(t) + Bw(t)$, $y(t) =$

"Facts are not science - as the dictionary is not literature." *Martin Henry Fischer*

	$Cx(t) + Dw(t)$ is given by $G(s) = C(sI - A)^{-1}B + D$.
$\langle v, w \rangle$	The inner product of two continuous-time signals v and w which is defined by $\langle v, w \rangle = \int_{-\infty}^{\infty} v^T w \, dt$.
$\|w\|_2$	The 2-norm of the continuous-time signal w which is defined by $\|w\|_2 = \sqrt{\langle w, w \rangle} = \sqrt{\int_{-\infty}^{\infty} w^T w dt}$.
$\mathcal{L}_2$	The set of continuous-time signals having finite 2-norm, which is known as the Lebesgue 2-space.
$\lambda_i(A)$	The i eigenvalues of A.
Re{ $\lambda_i(A)$}	The real part of the eigenvalues of A.
Asymptotic stability	A linear system $\mathcal{G}$ is said to be asymptotically stable if its output $y \in \mathcal{L}_2$ for any $w \in \mathcal{L}_2$. If Re{$\lambda_i(A)$} are in the left-hand-plane or equivalently if the real part of transfer function's poles are in the left-hand-plane then the system is stable.
$\mathcal{G}^H$	The adjoint of $\mathcal{G}$. The adjoint of a system having the state-space parameters {A, B, C, D} is a system parameterised by {– A^T, – C^T, B^T, D^T}.
$G^H(s)$	The adjoint (or Hermitian transpose) of the transfer function matrix $G(s)$.
$\Phi_{zz}(s)$	The spectral density matrix of the measurements z.
$\Delta(s)$	The spectral factor of $\Phi_{zz}(s)$ which satisfies $\Delta\Delta^H(s) = GQG^H(s) + R$ and $\Delta^{-H}(s) = (\Delta^H)^{-1}(s)$.
$G^{-1}(s)$	Inverse of the transfer function matrix $G(s)$.
$G^{-H}(s)$	Inverse of the adjoint transfer function matrix $G^H(s)$.
$\{G(s)\}_+$	Causal part of the transfer function matrix $G(s)$.
$H(s)$	Transfer function matrix of the minimum mean-square-error solution.
$H_{OE}(s)$	Transfer function matrix of the minimum mean-square-error solution specialised for output estimation.
$H_{IE}(s)$	Transfer function matrix of the minimum mean-square-error solution specialised for input estimation.

"Facts are stupid things." *Ronald Wilson Reagan*

1.7 References

[1] O. Neugebauer, *A history of ancient mathematical astronomy*, Springer, Berlin and New York, 1975.

[2] C. F. Gauss, *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientum*, Hamburg, 1809 (Translated: *Theory of the Motion of the Heavenly Bodies*, Dover, New York, 1963).

[3] A. N. Kolmogorov, "Sur l'interpolation et extrapolation des suites stationaires", *Comptes Rendus. de l'Academie des Sciences*, vol. 208, pp. 2043 – 2045, 1939.

[4] N. Wiener, *Extrapolation, interpolation and smoothing of stationary time series with engineering applications*, The MIT Press, Cambridge Mass.; Wiley, New York; Chapman & Hall, London, 1949.

[5] P. Masani, "Wiener's Contributions to Generalized Harmonic Analysis, Prediction Theory and Filter Theory", *Bulletin of the American Mathematical Society*, vol. 72, no. 1, pt. 2, pp. 73 – 125, 1966.

[6] T. Kailath, *Lectures on Wiener and Kalman Filtering*, Springer Verlag, Wien; New York, 1981.

[7] M.-A. Parseval Des Chênes, *Mémoires présentés à l'Institut des Sciences, Lettres et Arts, par divers savans, et lus dans ses assemblées. Sciences mathématiques et physiques (Savans étrangers)*, vol. 1, pp. 638 – 648, 1806.

[8] C. A. Desoer and M. Vidyasagar, *Feedback Systems : Input Output Properties*, Academic Press, N.Y., 1975.

[9] G. A. Einicke, "Asymptotic Optimality of the Minimum-Variance Fixed-Interval Smoother", *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1543 – 1547, Apr. 2007.

[10] T. Kailath, *Linear Systems*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1980.

[11] D. J. N. Limebeer, B. D. O. Anderson, P. Khargonekar and M. Green, "A Game Theoretic Approach to H_∞ Control for Time-varying Systems", *SIAM Journal of Control and Optimization*, vol. 30, no. 2, pp. 262 – 283, 1992.

[12] M. Green and D. J. N. Limebeer, *Linear Robust Control*, Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1995.

[13] C. S. Burrus, J. H. McClellan, A. V. Oppenheim, T. W. Parks, R. W. Schafer and H. W. Schuessler, *Computer-Based Exercises for Signal Processing Using Matlab*, Prentice-Hall, Englewood Cliffs, New Jersey, 1994.

[14] U. Shaked, "A general transfer function approach to linear stationary filtering and steady state optimal control problems", *International Journal of Control*, vol. 24, no. 6, pp. 741 – 770, 1976.

[15] A. H. Sayed and T. Kailath, "A Survey of Spectral Factorization Methods", *Numerical Linear Algebra with Applications*, vol. 8, pp. 467 – 496, 2001.

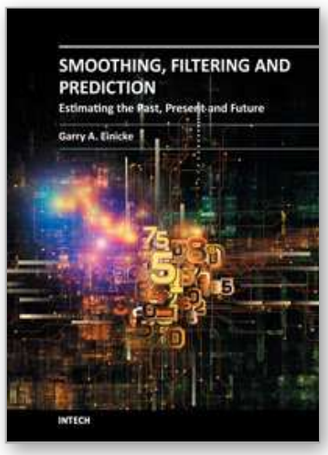
[16] U. Shaked, "H_∞–Minimum Error State Estimation of Linear Stationary Processes", *IEEE Transactions on Automatic Control*, vol. 35, no. 5, pp. 554 – 558, May 1990.

[17] S. A. Kassam and H. V. Poor, "Robust Techniques for Signal Processing: A Survey", *Proceedings of the IEEE*, vol. 73, no. 3, pp. 433 – 481, Mar. 1985.

[18] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*, McGraw-Hill Book Company, New York, 1971.

"All science is either physics or stamp collecting." *Baron William Thomson Kelvin*

IntechOpen



Smoothing, Filtering and Prediction - Estimating The Past, Present and Future

Edited by

ISBN 978-953-307-752-9
Hard cover, 276 pages
Publisher InTech
Published online 24, February, 2012
Published in print edition February, 2012

This book describes the classical smoothing, filtering and prediction techniques together with some more recently developed embellishments for improving performance within applications. It aims to present the subject in an accessible way, so that it can serve as a practical guide for undergraduates and newcomers to the field. The material is organised as a ten-lecture course. The foundations are laid in Chapters 1 and 2, which explain minimum-mean-square-error solution construction and asymptotic behaviour. Chapters 3 and 4 introduce continuous-time and discrete-time minimum-variance filtering. Generalisations for missing data, deterministic inputs, correlated noises, direct feedthrough terms, output estimation and equalisation are described. Chapter 5 simplifies the minimum-variance filtering results for steady-state problems. Observability, Riccati equation solution convergence, asymptotic stability and Wiener filter equivalence are discussed. Chapters 6 and 7 cover the subject of continuous-time and discrete-time smoothing. The main fixed-lag, fixed-point and fixed-interval smoother results are derived. It is shown that the minimum-variance fixed-interval smoother attains the best performance. Chapter 8 attends to parameter estimation. As the above-mentioned approaches all rely on knowledge of the underlying model parameters, maximum-likelihood techniques within expectation-maximisation algorithms for joint state and parameter estimation are described. Chapter 9 is concerned with robust techniques that accommodate uncertainties within problem specifications. An extra term within Riccati equations enables designers to trade-off average error and peak error performance. Chapter 10 rounds off the course by applying the afore-mentioned linear techniques to nonlinear estimation problems. It is demonstrated that step-wise linearisations can be used within predictors, filters and smoothers, albeit by forsaking optimal performance guarantees.

Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen

IntechOpen