# A Causal Optimal Filter of the Second Degree

Anatoli Torokhti*, Phil Howlett† and Charles Pearce‡

We provide a *non-linear* optimal physically realizable filter which guarantees a smaller associated error than those of the known *linear* optimal filters proposed in [1] and [2]. The technique is based on the best approximation of a stochastic signal by a specific non-linear operator formed from lower triangular matrices.

## 1 Introduction

Bode and Shannon [1] proposed a filter which has a smallest associated error in the class of physically realizable *linear* filters. Ruzhansky and Fomin [2] extended the result [1] to the so called 'weighted' *linear* filter. In this paper, we provide a *non-linear* physically realizable filter which guarantees a smaller associated error than those of the filters considered in [1], [2]. See Corollary 1 in this connection. Our approach is based on an extension of the techniques presented in [1] – [4]. Unlike the methodology provided in [1] – [4] and [6] – [12], the proposed filter satisfies the additional constraint (3) below.

Let $(\Omega, \Sigma, \mu)$ be a probability space, where $\Omega$ is the set of outcomes, $\Sigma$ a $\sigma$–field of measurable subsets of $\Omega$ and $\mu : \Sigma \mapsto [0, 1]$ an associated probability measure on $\Sigma$ with $\mu(\Omega) = 1$. Suppose that $x \in L^2(\Omega, \mathbb{R}^m)$ and $y \in L^2(\Omega, \mathbb{R}^m)$ are random vectors with realizations $x(\omega) \in \mathbb{R}^m$ and $y(\omega) \in \mathbb{R}^m$, respectively. Suppose $y(\omega)$ is observable data and $x(\omega)$ is an unobservable signal.

Each operator $F : \mathbb{R}^n \to \mathbb{R}^m$ defines an associated operator $\mathcal{F}_F : L^2(\Omega, \mathbb{R}^m) \to L^2(\Omega, \mathbb{R}^m)$ via the equation $[\mathcal{F}_F(y)](\omega) = F[y(\omega)]$ for each $\omega \in \Omega$. It is customary to write $F(y)$ rather than $\mathcal{F}_F(y)$, since we then have $[F(y)](\omega) = F[y(\omega)]$ for each $\omega \in \Omega$. It is also convenient to write $y$ instead of $y(\omega)$, $x$ instead of $x(\omega)$, etc.

Let $x = (x_1 \dots x_m)^T \in \mathbb{R}^m$ and $y = (y_1 \dots y_m)^T \in \mathbb{R}^m$. Each component $x_i$ (or $y_i$) can be interpreted as a value of $x$ (or $y$, respectively) at time $t_i$. We denote by $\hat{x}_i$ an estimate of $x_i$ for $i = 1, \dots, m$.

Let $F : \mathbb{R}^m \to \mathbb{R}^m$ be a filter defined by equations $\hat{x} = F(y)$ or

$$\hat{x}_i = f_{im}(y_1, \dots, y_m)$$

with $\hat{x} = (\hat{x}_1 \dots \hat{x}_m)^T$ and $f_{im} : \mathbb{R}^m \to \mathbb{R}$ for $i = 1, \dots, m$.

The filter $F$ is called physically realizable or causal if its estimate $\hat{x}_i$ of the signal component $x_i$ is determined from observable components $y_1, \dots, y_k$ of data $y$ with $k \le i$, i.e. if

$$\hat{x}_i = f_{ik}(y_1, \dots, y_k) \quad \text{with} \quad k \le i.$$

If $F$ is linear, i.e. $F$ is a matrix, then the latter condition means that $F$ is lower triangular.

We consider a class of non-linear filters $F$ given by

$$F(y) = F_0 + F_1 y + F_2 y^2, \tag{1}$$

where $F_0 \in \mathbb{R}^m$, $F_1, F_2 \in \mathbb{R}^{m \times m}$ and $y^2$ is determined by the Hadamard product so that $y^2 = (y_1^2 \dots y_m^2)^T$ with $y_1, \dots, y_m \in \mathbb{R}$.

*Centre for Industrial and Applicable Mathematics, University of South Australia, Mawson Lakes, SA 5095, Australia and Applied Mathematics Department, University of Adelaide, Adelaide, SA 5001, Australia. *Phone:* +61 8 8302 3812. *Fax:* +61 8 8302 5785 (*e-mail:* anatoli.torokhti@unisa.edu.au).

†Centre for Industrial and Applicable Mathematics, University of South Australia, Mawson Lakes, SA 5095, Australia. *Phone:* +61 8 8302 3195. *Fax:* +61 8 8302 5785 (*e-mail:* phil.howlett@unisa.edu.au).

‡Applied Mathematics Department, University of Adelaide, Adelaide, SA 5005, Australia. *Phone:* +61 8 8303 5412. *Fax:* +61 8 8303 3696 (*e-mail:* cpearce@maths.adelaide.edu.au)

Let $\mathcal{M}_+^{m\times m}$ be a set of $m \times m$ lower traingular matrices. We wish to find $F_0^0$, $F_1^0$, $F_2^0$ so that

$$J(F_0^0, F_1^0, F_2^0) = \min_{F_0, F_1, F_2} J(F_0, F_1, F_2) \qquad (2)$$

subject to

$$F_1 \in \mathcal{M}_+^{m\times m} \quad \text{and} \quad F_2 \in \mathcal{M}_+^{m\times m}, \qquad (3)$$

where

$$J(F_0, F_1, F_2) = E[\|x - F(y)\|^2] \qquad (4)$$

with $E$ the expectation operator and $\|\cdot\|$ the Frobenius norm.

The condition $F_1,\ F_2 \in \mathcal{M}_+^{m\times m}$ implies the causality of the filter satisfying (2).

Note, that a particular case of (1) – (4) with $F_0 = \mathbb{O}$ and $F_2 = \mathbb{O}$, where $\mathbb{O}$ is the zero vector or zero matrix, coincides with the problems considered in [1] and [2].

## 2 Solution of the Problem

Let $z = y^2$, $E_{xy} = E[xy^T] - E[x]E[y^T]$, $E_{yy} = E[yy^T] - E[y]E[y^T]$, $D = E_{zz} - E_{zy}E_{yy}^\dagger E_{yz}$ and $G = E_{xz} - E_{xy}E_{yy}^\dagger E_{yz}$, where $A^\dagger$ is the Moore-Penrose pseudo-inverse of $A \in \mathbb{R}^{m\times m}$.

We note, that matrix $A = \{a_{ij}\}$, where $a_{ij}$ is its entry for $i, j = 1, \ldots, m$, can always be represented in the form

$$A = A_+ + A_-,$$

where $A_+ = \{a_{ij}\}$ with $a_{ij} = 0$ for $i > j$, and $A_- = \{a_{ij}\}$ with $a_{ij} = 0$ for $i \le j$.

We denote $A^{-T} = (A^T)^{-1}$.

**Theorem 1** *Let $E_{yy}$ and $D$ be positive definite. Let*

$$E_{yy} = LL^T \quad \text{and} \quad D = MM^T$$

*be the Cholesky factorizations [5] for $E_{yy}$ and $D$, respectively.*[1] *Then*

$$F_0^0 = E[x] - F_1^0 E[y] - F_2^0 E[z], \qquad (5)$$

$$F_1^0 = ([E_{xy} - F_2^0 E_{zy}]L^{-T})_+ L^{-1} \in \mathcal{M}_+^{m\times m} \qquad (6)$$

*and*

$$F_2^0 = (GM^{-T})_+ M^{-1} \in \mathcal{M}_+^{m\times m}. \qquad (7)$$

[1] This means that $L, M \in \mathcal{M}_+^{m\times m}$.

**Proof** We have

$$J(F_0, F_1, F_2) = J_0 + J_1 + J_2 + J_3,$$

where

$$J_0 = \mathrm{tr}\{E_{xx} - E_{xy}E_{yy}^\dagger E_{yx}\} - \|G(D^\dagger)^{\frac{1}{2}}\|^2, \qquad (8)$$

$$J_1 = \|F_0 - (E[x] - F_1 E[y] - F_2 E[z])\|^2, \qquad (9)$$

$$J_2 = \|[F_1 - (E_{xy} - F_2 E_{zy})E_{yy}^\dagger]E_{yy}^{1/2}\|^2 \qquad (10)$$

and

$$J_3 = \|[F_2 - GD^\dagger]D^{1/2}\|^2. \qquad (11)$$

See [3] and [4] in this regard.

Since $E_{yy}$ and $D$ are positive definite, we can write $E_{yy}^\dagger = E_{yy}^{-1}$ and $D^\dagger = D^{-1}$.

Next, let us denote

$$H_+ = (GM^{-T})_+, \quad \text{and} \quad H_- = (GM^{-T})_-.$$

Then

$$\begin{aligned} J_3 &= \mathrm{tr}(F_2 M - H_+ - H_-)(M^T F_2^T - H_+^T - H_-^T) \\ &= \mathrm{tr}(F_2 M - H_+)(M^T F_2^T - H_+^T) \\ &\quad -\mathrm{tr}(F_2 M H_-^T + H_- M^\dagger F_2^T) \\ &\quad +\mathrm{tr}(H_+ H_-^T + H_- H_+^T + H_- H_-^T) \\ &= \mathrm{tr}(F_2 M - H_+)(M^T F_2^T - H_+^T). \end{aligned}$$

Here

$$\mathrm{tr}(H_+ H_-^T + H_- H_+^T + H_- H_-^T) = 0$$

and

$$\mathrm{tr}(F_2 M H_-^T + H_- M^\dagger F_2^T) = 0$$

because we consider lower triangular matrices $F_2$ only. Therefore

$$F_2^0 = H_+ M^{-1} = (GM^{-T})_+ M^{-1} \in \mathcal{M}_+^{m\times m}.$$

Similarly, if we denote

$$K_+ = ([E_{xy} - F_2 E_{zy}]L^{-T})_+$$

and

$$K_- = ([E_{xy} - F_2 E_{zy}]L^{-T})_-$$

then

$$\begin{aligned} J_2 &= \mathrm{tr}(F_1 L - K_+)(L^T F_1 - K_+^T) \\ &\quad -\mathrm{tr}(F_1 L K_-^T + K_- L^T F_1^T) \\ &\quad +\mathrm{tr}(K_+ K_-^T + K_- K_+^T + K_- K_-^T) \\ &= \mathrm{tr}(F_1 L - K_+)(L^T F_1 - K_+^T) = \|F_1 L - K_+\|^2, \end{aligned}$$

which implies $F_1 = F_1^0$.

The equation $F_0 = F_0^0$ follows directly from (8). □

**Theorem 2** *The error associated with the optimal filter $F^0$ defined by the equation*

$$F^0(y) = F_0^0 + F_1^0 y + F_2^0 z$$

*is*

$$E[||x - F^0(y)||^2]$$
$$= tr\{E_{xx} - E_{xy}E^{-1}E_{yx}\} - ||G(D^{-1})^{\frac{1}{2}}||^2. \quad (12)$$

**Proof** The proof follows from (8) – (10) with $F_0 = F_0^0$, $F_1 = F_1^0$ and $F_2 = F_2^0$. □

## 3 Discussion

The Bode-Shannon filter $B^0$ [1], [2] is determined by the equation

$$B^0 = (E[xy^T]L^{-T})_+ L^{-1}.$$

**Theorem 3** *The error associated with the Bode-Shannon filter is*

$$E[||x - B^0(y)||^2] = tr\{E_{xx} - E_{xy}E^{-1}E_{yx}\}. \quad (13)$$

**Corollary 1** *The error associated with the filter $F^0$ is less than that of the filter $B^0$ for $||G(D^{-1})^{\frac{1}{2}}||^2$, i.e.*

$$E[||x - B^0(y)||^2] - E[||x - F^0(y)||^2] = ||G(D^{-1})^{\frac{1}{2}}||^2.$$

**Proof** The proofs of Theorem 3 and Corollary 1 follow from the above. □

Next, Ruzhansky and Fomin [2] provided a filter $R^0$ which satisfies the condition

$$\mathcal{J}(R^0) = \min_R \mathcal{J}(R) \quad (14)$$

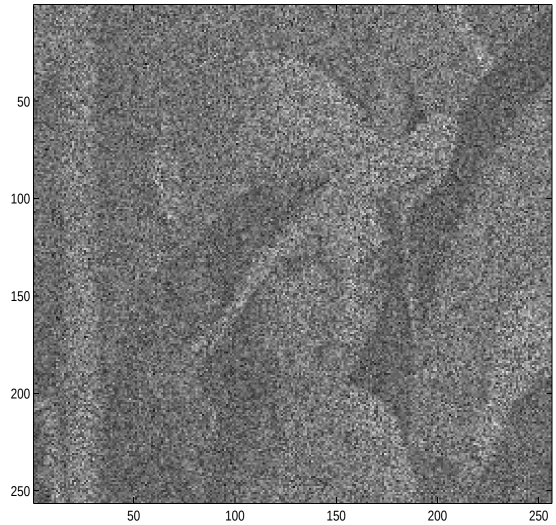
over all lower triangular matrices $R$, where

$$\mathcal{J}(R) = E[||W(x - Ry)||^2]$$

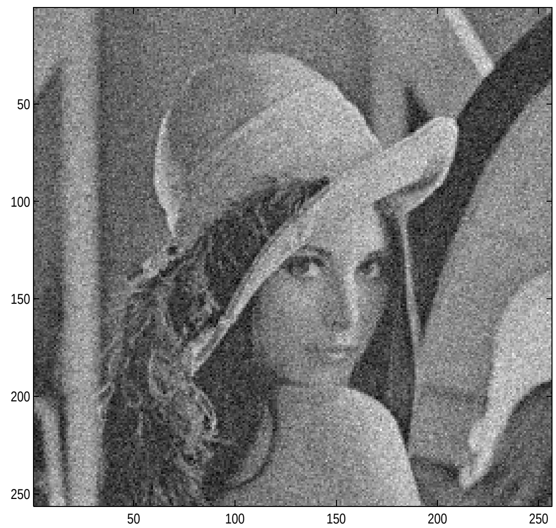
and $W$ is a weight matrix. If we denote $u = Wx$ and $P = WR$ then the problem (14) becomes a particular case of the problem (1) – (4) with $F_0 = \mathbb{O}$, $F_1 = P$ and $F_2 = \mathbb{O}$.

Therefore the error associated with the filter $R^0$ coincides with (13) if we replace $x$ by $u = Wx$ and $B^0$ by $R^0$. As a result, Corollary 1 is true for the filter $R^0$ as well if change notation.

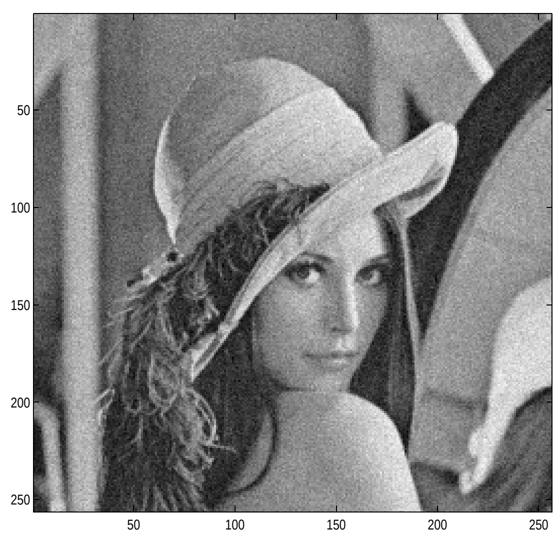
Corollary 1 demonstrates the advantages of our approach.




(a) Oserved data.




(b) Esimate $X_B$ by $B^0$.




(c) Esimate $X_F$ by $F^0$.

## 4 Simulation

To illustrate the performanve of the filter $F^0$, we applied the proposed technique to the filtering of the known image "Lena" corrupted by a combination of additive and multiplicative noise. The observed noisy data has been presented by a matrix $Y \in \mathbb{R}^{256\times256}$ in the form

$$Y = 250N_1 + 30XN_2,$$

where $N_1$ is a matrix with normally distributed entries with mean 0 and variance 1, $N_2$ is a matrix with uniformly distributed entries in the interval $(0,1)$, and $X \in \mathbb{R}^{256\times256}$ is a numerical representation of the image. The simulation demonstrates a clear advantage of the filter $F^0$ over the filter $B^0$. For $X_F$ and $X_B$ to be the estimates by $F^0$ and $B^0$, respectively, the relations of the errors are

$$||X - X_B||^2 - ||X - X_F||^2 = 15.2401 \times 10^6$$

and $||X - X_B||^2/||X - X_F||^2 = 2.4$.

## 5 Final Remark

We have presented a new technique allowing us to find the optimal non-linear filter $F^0$ which guarantees better accuracy compared with that of the known optimal linear filters considered in [1] and [2]. The important feature of our approach is physical realizability of the provided filter. The clear superiority of the filter $F^0$ over filters [1] and [2] has been justified.

Potential applications of the proposed technique are abundant including, for example, image processing [3], [4]; data compression [6]; some areas in pattern recognition; blind channel equalization; target detection; optimal nonlinear system synthesis [4], [7], [9] [10], etc. The authors shall focus further work on extensions of the proposed method to the problems considered in [4], [6] – [12].

## References


[1] H.W. Bode and C.E. Shannon, A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory, *Proc. IRE,* vol. 38, pp. 417–425, 1950.

[2] M.V. Ruzhansky and V.N. Fomin, Optimal Filter Construction for a General Quadratic Cost Functional, *Bulletin of St. Petersburg University. Mathematics,* vol. 28, pp. 50–55, 1995.

[3] A. Torokhti and P. Howlett, 'An Optimal Filter of the Second Order', *IEEE Trans. on Signal Processing*, vol. 49, pp. 1044–1048, 2001.

[4] A. Torokhti and P. Howlett, 'On the Best Quadratic Approximation of Nonlinear Systems', *IEEE Trans. on Circuits and Systems. Part I, Fundamental theory and applications*, vol. 48, pp. 595–602, 2001.

[5] G.H Golub and C.F. Van Loan, *Matrix Computations,* Baltimore, Johns Hopkins University Press, 1996.

[6] A. Torokhti and P. Howlett, 'Optimal Fixed Rank Transform of the Second Degree', *IEEE Trans. on Circuits and Systems. Part II, Analog & Digital Signal Processing* vol. 48, pp. 309–315, 2001.

[7] P.G. Howlett, A.P. Torokhti, C.E.M. Pearce, 'The modelling and numerical simulation of causal non-linear systems', *Nonlinear Analysis: Theory, Methods & Applications,* vol. 47 (8), pp. 5559-5572, 2001.

[8] P.G. Howlett, C.E.M. Pearce, A. P. Torokhti, 'A best linear estimator for random vectors with values in Hilbert space', *Maths. Res.* (St. Petersburg) vol. 4, pp. 99-107, 1999.

[9] A. Torokhti, P. Howlett and C. Pearce, 'Method of Best Successive Approximations for Nonlinear Operators', *J. of Computational Analysis and Applications* (to appear).

[10] P. Howlett, A. Torokhti, The Methodology of Numerical Representation of Non-Linear Operators on Non-Compact Sets, *Numerical Functional Analysis and Optimization,* vol. 18, no. 3/4, pp. 343 - 366, 1997.

[11] P. Howlett, C. Pearce and A. Torokhti, 'On Problems of Best Approximation for Nonlinear Operators', *Nonlinear Functional Analysis and Applications* (to appear).

[12] P. Howlett, C. Pearce and A. Torokhti, 'An optimal linear filter for random signals with realisations in Hilbert Space', *ANZIAM Journal* (to appear).