CONVOLUTIVE REDUCED RANK WIENER FILTERING

Jonathan H. Manton and Yingbo Hua

ARC Special Research Centre for Ultra-Broadband Information Networks
Department of Electrical and Electronic Engineering
The University of Melbourne, Parkville, Victoria 3010, Australia.
j.manton@ee.unimelb.edu.au

ABSTRACT

If two wide sense stationary time series are correlated then one can be used to predict the other. The reduced rank Wiener filter is the rank constrained linear operator which maps the current value of one time series to an estimate of the current value of the other time series in an optimal way. A closed form solution exists for the reduced rank Wiener filter. This paper studies the problem of determining the reduced rank FIR filter which optimally predicts one time series given the other. This optimal FIR filter is called the convolutive reduced rank Wiener filter, and it is proved that determining it is equivalent to solving a weighted low rank approximation problem. In certain cases a closed form solution exists, and in general, the iterative optimisation algorithm derived here can be used to converge to a locally optimal convolutive reduced rank Wiener filter.

1. INTRODUCTION

Let $\boldsymbol{x}(t) \in \mathbb{R}^m$ and $\boldsymbol{y}(t) \in \mathbb{R}^n$ be two wide sense stationary time series, where $t \in \mathbb{R}$ denotes time. Let z^{-1} denote the unit delay operator defined by $z^{-1}\boldsymbol{x}(t) = \boldsymbol{x}(t-1)$. The convolutive rank r Wiener filter of order p is defined to be the matrix $T(z^{-1})$ satisfying

$$\arg \min_{\substack{T(z^{-1}) \\ T(z^{-1})=A(z^{-1})B}} \boldsymbol{E}\left[\left\|\boldsymbol{y}(t) - T(z^{-1})\boldsymbol{x}(t)\right\|^2\right],$$

$$A(z^{-1}) = \sum_{i=0}^{p-1} A_i z^{-i}, \quad A_i \in \mathbb{R}^{n\times r}, \quad B \in \mathbb{R}^{r\times m} \qquad (1)$$

where the minimisation is over the elements of the matrices A_i and B. (Since $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ are wide sense stationary, the expectation in (1) does not depend on t.) The norm is taken to be the 2-norm; $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x}$ for any vector $\boldsymbol{x}$, where the superscript T denotes transpose. The special case of $p = 1$ in (1) has been extensively studied in the literature; the minimising matrix T is called the reduced rank Wiener filter [2, 3, 4, 9, 10, 11, 12]. This paper studies various properties of the convolutive reduced rank Wiener filter, including deriving a closed form solution of (1) under certain conditions on the statistics of $\boldsymbol{x}(t)$.

The apparently more general case of when B is allowed to be a polynomial in z^{-1} reduces to (1) as follows. If $B(z^{-1}) = \sum_{i=0}^{q-1} B_i z^{-i}$ then form the augmented matrix $\tilde{B}$ and augmented vector $\tilde{\boldsymbol{x}}(t)$ given by

$$\tilde{B} = [B_0 \; B_1 \; \cdots \; B_{q-1}], \qquad (2)$$

$$\tilde{\boldsymbol{x}}(t) = \left[\boldsymbol{x}(t)^T \; \boldsymbol{x}(t-1)^T \; \cdots \; \boldsymbol{x}(t-q+1)^T\right]^T. \qquad (3)$$

Because $B(z^{-1})\boldsymbol{x}(t) = \tilde{B}\tilde{\boldsymbol{x}}(t)$, the $T(z^{-1}) = A(z^{-1})B(z^{-1})$ which minimises $\boldsymbol{E}\left[\|\boldsymbol{y}(t) - T(z^{-1})\boldsymbol{x}(t)\|^2\right]$ is found by solving (1) with $\boldsymbol{x}(t)$ replaced by $\tilde{\boldsymbol{x}}(t)$.

Motivation for considering (1) is now given. The matrix B can be thought of as an *analysis filter*. It extracts from the time series $\boldsymbol{x}(t)$ the r most useful linear combinations of the elements of $\boldsymbol{x}(t)$ for predicting $\boldsymbol{y}(t)$. The matrix $A(z^{-1})$ can be thought of as a *synthesis filter*. It optimally combines the r components of $B\boldsymbol{x}(t)$ to estimate $\boldsymbol{y}(t)$. Purposely choosing $r < m$ makes the resulting filter robust to mis-specification of the statistics of $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$. Indeed, restricting B to have r rows is similar in spirit to principal component analysis [1]. Furthermore, rank reduction is known to trade bias for variance (or risk) [7, 11].

The outline of this paper is as follows. Section 2 briefly reviews the weighted low rank approximation problem. Section 3 proves the main result of this paper, which is that the convolutive rank reduced Wiener filter can be calculated by solving an associated weighted low rank approximation problem. This fundamental result is used in Section 4 to derive a closed form solution of (1) under certain conditions on the statistics of $\boldsymbol{x}(t)$. A numerical algorithm for solving (1) is given in Section 5. Simulations in Section 6 confirm that the convolutive rank reduced Wiener filter performs as expected.

2. LOW RANK APPROXIMATION

This section briefly reviews the weighted low rank approximation problem [8] of calculating

$$\arg \min_{\substack{R \\ \text{rank}\{R\}\le r}} \|X - R\|_Q^2,$$

$$\|X - R\|_Q^2 = \text{vec}\{X - R\}^T Q \,\text{vec}\{X - R\} \qquad (4)$$

for a given data matrix $X \in \mathbb{R}^{np\times m}$ and positive definite symmetric weighting matrix $Q \in \mathbb{R}^{mnp\times mnp}$. Here, $\text{vec}\{\cdot\}$ is the vec operator [5] which stacks the columns of a matrix to form a column vector. Section 3 will show that the convolutive reduced rank Wiener filter problem (1) can be solved by solving a corresponding low rank approximation problem (4).

The traditional approach to reduced rank problems [12] is to over-parameterise R as $R = AB$ where A has r columns and

This work was supported by the Australian Research Council.

B has r rows. However, some properties of the reduced rank problem, such as whether or not a closed form solution exists, are obscured by the ambiguity in the decomposition $R = AB = \left(AS^{-1}\right)\left(SB\right)$, the latter equality holding for any invertible matrix $S \in \mathbb{R}^{r\times r}$.

The key insight [8] required to remove the ambiguity is the realisation that the matrix R can be decomposed as $R = AB$ if and only if there exists a full rank matrix $N \in \mathbb{R}^{m\times(m-r)}$ such that $RN = 0$. (Given either B or N, the other may be taken to be any matrix satisfying $BN = 0$.) This implies that (4) can be solved by first computing

$$\arg \min_{\substack{N\in\mathbb{R}^{m\times(m-r)}\\ \left|N^TN\right|\neq 0}} f_1(N), \qquad f_1(N) = \min_{\substack{R\in\mathbb{R}^{np\times m}\\ RN=0}} \|X-R\|_Q^2 \tag{5}$$

where $\left|N^TN\right|$ is the determinant of N^TN and is non-zero if and only if N has full rank. A closed form expression for $f_1(N)$ in (5) exists and is given by [8, Th. 1]

$$\begin{aligned} f_1(N) &= \operatorname{vec}\{X\}^T (N\otimes I_{np}) \\ &\left[(N\otimes I_{np})^T Q^{-1} (N\otimes I_{np})\right]^{-1} (N\otimes I_{np})^T \operatorname{vec}\{X\} \end{aligned} \tag{6}$$

where $\otimes$ is Kronecker's product [5] and $I_{np} \in \mathbb{R}^{np\times np}$ is the identity matrix.

The reformulation (5) has removed the ambiguity in the following sense. Since $f_1(NS) = f_1(N)$ for any invertible matrix S, the value of $f_1(N)$ depends only on the range space of N. Mathematically, if $G_{m,m-r}$ denotes the collection of all $m-r$ dimensional subspaces of $\mathbb{R}^m$ (so that $\operatorname{range}\{N\} \in G_{m,m-r}$ if N has full rank), then there exists a function $\tilde{f}_1 : G_{m,m-r} \to \mathbb{R}$ such that $f_1(N) = \tilde{f}_1(\operatorname{range}\{N\})$ holds for all full rank matrices $N \in \mathbb{R}^{m\times(m-r)}$. Thus (5) can be written as

$$\arg \min_{S\in G_{m,m-r}} \tilde{f}_1(S) \tag{7}$$

which is an unconstrained optimisation problem on the Grassmannian manifold $G_{m,m-r}$. Since (7) has a unique solution in general, the ambiguity has been removed [8].

Although (7) will not be used subsequently in this paper, the reason for mentioning it is that it shows that all the information about the rank reduced problem (4) is nicely captured by the reformulation (5). The same idea is used in the next section to study the convolutive reduced rank Wiener filter.

3. RELATIONSHIP TO LOW RANK APPROXIMATION

This section shows that the convolutive reduced rank Wiener filter (1) can be computed by solving a weighted low rank approximation problem (4).

Finding a closed form solution to the convolutive reduced rank Wiener filter is impeded by the ambiguity in the decomposition $T(z^{-1}) = A(z^{-1})B = \left(A(z^{-1})S^{-1}\right)\left(SB\right)$, the latter equality holding for any invertible matrix $S \in \mathbb{R}^{r\times r}$. As in Section 2, this ambiguity is removed by noting that the matrix $T(z^{-1}) = T_0 + T_1z^{-1} + \cdots + T_{p-1}z^{-(p-1)}$ can be decomposed as $T(z^{-1}) = A(z^{-1})B$ if and only if there exists a full rank matrix N such that $T_0N = T_1N = \cdots = T_{p-1}N = 0$. (Given either B or N, the other may be taken to be any matrix satisfying $BN = 0$.) Thus (1) can be solved by first computing

$$\arg \min_{\substack{N\in\mathbb{R}^{m\times(m-r)}\\ \left|N^TN\right|\neq 0}} f_2(N),$$
$$f_2(N) = \min_{\substack{\tilde{T}\in\mathbb{R}^{n\times pm}\\ \tilde{T}(I_p\otimes N)=0}} \boldsymbol{E}\left[\|\boldsymbol{y}(t) - \tilde{T}\tilde{\boldsymbol{x}}(t)\|^2\right] \tag{8}$$

where

$$\begin{aligned} \tilde{T} &= [T_0\; T_1\; \cdots\; T_{p-1}], \\ \tilde{\boldsymbol{x}}(t) &= \left[\boldsymbol{x}(t)^T\; \boldsymbol{x}(t-1)^T\; \cdots\; \boldsymbol{x}(t-p+1)^T\right]^T \end{aligned} \tag{9}$$

are the augmented versions of the components of $T(z^{-1})$ and $\boldsymbol{x}(t)$. (Note that $\tilde{T}(I_p\otimes N) = [T_0N\; T_1N\; \cdots\; T_{p-1}N]$.)

It is assumed that the following covariance matrices are given.

$$R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} = \boldsymbol{E}\left[\tilde{\boldsymbol{x}}(t)\tilde{\boldsymbol{x}}(t)^T\right], \quad R_{\boldsymbol{y}\tilde{\boldsymbol{x}}} = \boldsymbol{E}\left[\boldsymbol{y}(t)\tilde{\boldsymbol{x}}(t)^T\right], \tag{10}$$
$$R_{\boldsymbol{y}\boldsymbol{y}} = \boldsymbol{E}\left[\boldsymbol{y}(t)\boldsymbol{y}(t)^T\right], \quad R_{\tilde{\boldsymbol{x}}\boldsymbol{y}} = \boldsymbol{E}\left[\tilde{\boldsymbol{x}}(t)\boldsymbol{y}(t)^T\right]. \tag{11}$$

Since $\boldsymbol{E}\left[\|\boldsymbol{y}(t) - \tilde{T}\tilde{\boldsymbol{x}}(t)\|^2\right]$ is quadratic in the elements of $\tilde{T}$, it is possible to choose X, R and Q in (5) so that $f_1(N) = f_2(N)+c$ for some constant c. This equivalence is now established formally.

To make the constraint $\tilde{T}(I_p\otimes N) = 0$ equivalent to $RN = 0$, define R to be $R = [T_0^T\; \cdots\; T_{p-1}^T]^T$. The cost can be written in terms of R instead of $\tilde{T}$ as follows. Define K to be the unique permutation matrix such that $K\operatorname{vec}\{R\} = \operatorname{vec}\left\{\tilde{T}^T\right\}$ holds for any choice of $T_0, \cdots, T_{p-1}$. (In fact, $K = (K_{n,p}\otimes I_m)K_{np,m}$ where the commutation matrix [5] $K_{np,m}$ is the unique permutation matrix such that $K_{np,m}\operatorname{vec}\{X\} = \operatorname{vec}\{X^T\}$ holds for any matrix $X \in \mathbb{R}^{np\times m}$, and similarly for $K_{n,p}$.) Define $c = \operatorname{tr}\left\{R_{\boldsymbol{y}\boldsymbol{y}} - R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}^{-1}R_{\tilde{\boldsymbol{x}}\boldsymbol{y}}\right\}$ and $\tilde{X} = R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}^{-1}$. Then

$$\begin{aligned} &\boldsymbol{E}\left[\|\boldsymbol{y}(t) - \tilde{T}\tilde{\boldsymbol{x}}(t)\|^2\right] = \operatorname{tr}\left\{R_{\boldsymbol{y}\boldsymbol{y}} - 2\tilde{T}R_{\tilde{\boldsymbol{x}}\boldsymbol{y}} + \tilde{T}R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}\tilde{T}^T\right\} \\ &= \operatorname{tr}\left\{(\tilde{X}-\tilde{T})R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}(\tilde{X}-\tilde{T})^T\right\} + c \\ &= \operatorname{vec}\left\{\tilde{X}^T - \tilde{T}^T\right\}^T (I_n\otimes R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}})\operatorname{vec}\left\{\tilde{X}^T - \tilde{T}^T\right\} + c \\ &= \operatorname{vec}\{X-R\}^T K^T(I_n\otimes R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}})K\operatorname{vec}\{X-R\} + c \end{aligned} \tag{12}$$

where X satisfies $K\operatorname{vec}\{X\} = \operatorname{vec}\left\{\tilde{X}^T\right\}$. This is summarised in Theorem 1 below.

Theorem 1 *The convolutive rank r Wiener filter $T(z^{-1}) = T_0 + T_1z^{-1} + \cdots + T_{p-1}z^{-(p-1)}$, when written in the form $R = [T_0^T\; \cdots\; T_{p-1}^T]^T$, is also the weighted low rank approximation of the matrix $X \in \mathbb{R}^{np\times m}$ with weighting $Q \in \mathbb{R}^{nmp\times nmp}$ where X and Q are given by*

$$\operatorname{vec}\{X\} = K^T\operatorname{vec}\left\{(R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}^{-1})^T\right\}, \tag{13}$$
$$Q = K^T(I_n\otimes R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}})K, \tag{14}$$
$$K = (K_{n,p}\otimes I_m)K_{np,m} \tag{15}$$

and the matrices $K_{n,p}$ and $K_{np,m}$ are the commutation matrices defined earlier.

Remark: Since $\mathrm{vec}\{R\} = K^T \mathrm{vec}\left\{\tilde{T}^T\right\}$, the matrix X is obtained from $R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R^{-1}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$ in the same way that R is obtained from $\tilde{T}$; divide $R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R^{-1}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$ into p blocks of size $n \times m$, then stack these blocks vertically.

From (12) it is readily seen that the optimal Wiener filter without the rank constraint is $R = X$, or equivalently, $\tilde{T} = R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R^{-1}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$. Therefore, Theorem 1 shows that the convolutive reduced rank Wiener filter is obtained by approximating a permuted version X of the convolutive full rank Wiener filter matrix $R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R^{-1}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$ by a rank reduced one. Somewhat surprisingly, the weighting matrix Q for the approximation does not depend on $R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}$ or $R_{\boldsymbol{y}\boldsymbol{y}}$.

It is remarked that the term $\mathrm{tr}\left\{R_{\boldsymbol{y}\boldsymbol{y}} - R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R^{-1}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}R_{\tilde{\boldsymbol{x}}\boldsymbol{y}}\right\}$ in (12) is the smallest mean square error (MSE) achievable if there is no rank constraint.

4. CLOSED FORM SOLUTION

For any matrix X with SVD $X = U\Sigma V^T$, define Σ_r to be Σ with all but the first r singular values set to zero. It is well known that the best rank r approximation of X in the unweighted case is $R = \mathrm{Trunc}_r\{X\} = U\Sigma_r V^T$. Theorem 2 below generalises this result to certain weighted cases. Combining Theorem 2 and Theorem 1 leads to a closed form solution of the convolutive rank reduced Wiener filter under certain conditions.

Theorem 2 *If Q in (4) can be decomposed as $Q = A \otimes B$, where $A \in \mathbb{R}^{m\times m}$ and $B \in \mathbb{R}^{np\times np}$, then the best rank r approximation of X is $R = B^{-\frac{1}{2}}\,\mathrm{Trunc}_r\left\{B^{\frac{1}{2}}XA^{\frac{1}{2}}\right\}A^{-\frac{1}{2}}$.*

PROOF. Define $\tilde{X} = B^{\frac{1}{2}}XA^{\frac{1}{2}}$ and $\tilde{N} = A^{-\frac{1}{2}}N$. Then $f_1(N)$, defined in (6), can be written as

$$f_1(N) = \mathrm{tr}\left\{\tilde{N}^T\tilde{X}^T\tilde{X}\tilde{N}\left(\tilde{N}^T\tilde{N}\right)^{-1}\right\}.$$

This generalised Rayleigh quotient [8] achieves its minimum when the columns of $\tilde{N}$ are the $m-r$ smallest right singular vectors of $\tilde{X}$. The R which minimises $\|X - R\|^2_Q$ in (5) subject to $R\tilde{N} = RA^{\frac{1}{2}}N = 0$ can then be shown to be (c.f., [8]) as given in the theorem. □

The following two results follow straightforwardly.

Lemma 3 *Define Q as in Theorem 1. Then Q can be decomposed as $Q = A \otimes B$, where $A \in \mathbb{R}^{m\times m}$ and $B \in \mathbb{R}^{np\times np}$, if and only if $R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$ can be decomposed as $R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} = R^{(1)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} \otimes R^{(2)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$, where $R^{(1)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} \in \mathbb{R}^{p\times p}$ and $R^{(2)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} \in \mathbb{R}^{m\times m}$.*

Corollary 4 (Optimal Filter) *Define X and R as in Theorem 1. If $R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$ decomposes as $R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} = R^{(1)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} \otimes R^{(2)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$, where $R^{(1)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} \in \mathbb{R}^{p\times p}$ and $R^{(2)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}} \in \mathbb{R}^{m\times m}$, then the convolutive rank reduced Wiener filter R is given by*

$$R = \left[\left(R^{(1)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}\right)^{-\frac{1}{2}} \otimes I_n\right] \mathrm{Trunc}_r\left\{\left[\left(R^{(1)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}\right)^{\frac{1}{2}} \otimes I_n\right] X \left(R^{(2)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}\right)^{\frac{1}{2}}\right\}\left(R^{(2)}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}\right)^{-\frac{1}{2}}. \quad (16)$$

If $p = 1$ then (16) reduces to

$$R = \mathrm{Trunc}_r\left\{R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R^{-\frac{1}{2}}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}\right\}R^{-\frac{1}{2}}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}},$$

the well-known formula for the rank reduced Wiener filter.

5. NUMERICAL ALGORITHM

This section presents a numerical algorithm for solving the convolutive rank reduced Wiener filtering problem (1). The algorithm is derived by first using Theorem 1 to convert it to a low rank approximation problem of the form (4), and then using one of the algorithms in [8] to solve (4).

Algorithm 5 *Given the correlation matrices $R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$ and $R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}$ defined in (10), the following algorithm iteratively converges to a local minimum of the cost function (1). The filter $T(z^{-1})$ is returned in the matrix form $R = [T_0^T \ \cdots \ T_{p-1}^T]^T$ where $T(z^{-1}) = T_0 + z^{-1}T_1 + \cdots + z^{-(p-1)}T_{p-1}$. The algorithm uses the function*

$$f(N) = \mathrm{vec}\{XN\}^T \left[(N \otimes I_n)^T Q^{-1}(N \otimes I_n)\right]^{-1} \mathrm{vec}\{XN\}. \quad (17)$$

1. *Set $K := (K_{n,p} \otimes I_m)K_{np,m}$. (See (1) for the definition of n, m and p, and see Section 3 for the definition of the commutation matrices $K_{np,m}$ and $K_{n,p}$.)*
2. *Set $Q := K^T(I_n \otimes R_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}})K$. Set $X \in \mathbb{R}^{np\times m}$ to the matrix for which $\mathrm{vec}\{X\} = K^T \mathrm{vec}\left\{(R_{\boldsymbol{y}\tilde{\boldsymbol{x}}}R^{-1}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}})^T\right\}$.*
3. *Set step size $\lambda := 1$. Choose $N \in \mathbb{R}^{m\times(m-r)}$ and $N_\perp \in \mathbb{R}^{m\times r}$ such that $[N\ N_\perp]^T[N\ N_\perp] = I$.*
4. *Set $A \in \mathbb{R}^{np\times(m-r)}$ and $B \in \mathbb{R}^{np\times m}$ to the matrices for which*
$$\mathrm{vec}\{A\} = \left[(N \otimes I_n)^T Q^{-1}(N \otimes I_n)\right]^{-1} \mathrm{vec}\{XN\},$$
$$\mathrm{vec}\{B\} = Q^{-1}\,\mathrm{vec}\left\{AN^T\right\}.$$
5. *Compute the descent direction $K = -2N_\perp^T(X - B)^T A$. If $\|K\|$ is sufficiently small then stop, returning the matrix $R \in \mathbb{R}^{np\times m}$ satisfying*
$$\mathrm{vec}\{R\} = \mathrm{vec}\{X\} - Q^{-1}(N \otimes I_n) \left[(N \otimes I_n)^T Q^{-1}(N \otimes I_n)\right]^{-1}(N \otimes I_n)^T \mathrm{vec}\{X\}.$$
6. *If $f(N) - f(N + 2\lambda N_\perp K) \geq \lambda\|K\|^2$ then set $\lambda := 2\lambda$ and repeat Step 6. (Recall that f is defined in (17).)*
7. *If $f(N) - f(N + \lambda N_\perp K) < \frac{1}{2}\lambda\|K\|^2$ then set $\lambda := \frac{1}{2}\lambda$ and repeat Step 7.*
8. *Set $N := N + \lambda N_\perp K$. Renormalise $[N\ N_\perp]$ by setting $[N\ N_\perp] := \mathrm{qf}\{N\}$. (The "Q-Factor" operator $\mathrm{qf}\{N\}$ is the Q-matrix in the QR decomposition of N.) Go to Step 4.*

Remark: Alg. 5 minimises (8) on the Grassmann manifold of matrices of the form $\left\{N : N^T N = I\right\}$. See [8, 6] for details.

6. SIMULATIONS

The following model was used to generate the time series $\boldsymbol{y}(t) \in \mathbb{R}^3$.

$$\boldsymbol{y}(t) = A_0\boldsymbol{x}(t) + A_1\boldsymbol{x}(t-1) + \sigma^2\boldsymbol{n}(t) \tag{18}$$

where both $\boldsymbol{x}(t)$ and $\boldsymbol{n}(t)$ are independent white Gaussian noise processes with zero mean and unit variance and the matrices A_0 and A_1 are given by

$$A_0 = \begin{bmatrix} 1 & 0.5 & 0.2 \\ 0.7 & 1 & 0.3 \\ 0.5 & 0.5 & 0.1 \end{bmatrix}, \qquad A_1 = \begin{bmatrix} 0.1 & 0.9 & 0.1 \\ 0.8 & 0.2 & 0.2 \\ 0.4 & 0.6 & 0.1 \end{bmatrix}. \tag{19}$$

Note that both A_0 and A_1 have full rank. In the figures, the noise variance σ^2 is related to the SNR according to the formula $\mathrm{SNR} = 10\log_{10}\sigma^2$.

The time series $\boldsymbol{x}(t)$ was used to predict $\boldsymbol{y}(t)$ using three filters; the Wiener filter $\widehat{\boldsymbol{y}}(t) = A_0\boldsymbol{x}(t)$, the convolutive rank 2 Wiener filter found by solving (1) with $r = 2$, and the convolutive (full rank) Wiener filter $\widehat{\boldsymbol{y}}(t) = A_0\boldsymbol{x}(t) + A_1\boldsymbol{x}(t-1)$. The results are plotted in Figure 1.

The same three filters were then used to predict $\boldsymbol{y}(t)$ given the noise corrupted time series $\boldsymbol{x}(t) + \boldsymbol{w}(t)$, where the additive white Gaussian noise $\boldsymbol{w}(t)$ had zero mean and variance 0.04. The results are plotted in Figure 2 and demonstrate the robustness of rank reduction to model mis-specification. The rank reduced convolutive Wiener filter performs similarly to the optimal full rank convolutive Wiener filter yet is computationally simpler to compute.

7. CONCLUSION

This paper introduced the convolutive rank reduced Wiener filter, which is a generalisation of the rank reduced Wiener filter previously studied in the literature. A closed form solution of the convolutive rank reduced Wiener filter was derived, as well as a numerical algorithm for computing it.

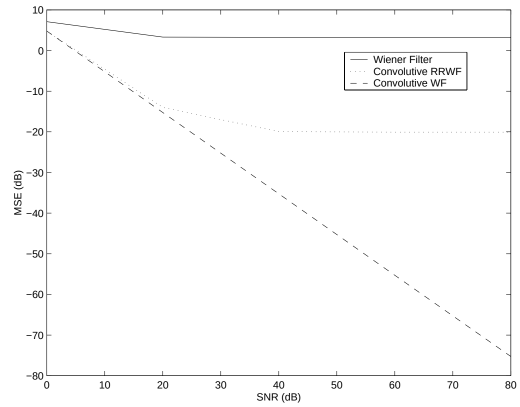

Figure 1: Graph comparing the performance of the Wiener filter, the convolutive (full rank) Wiener filter and the convolutive rank reduced Wiener filter.

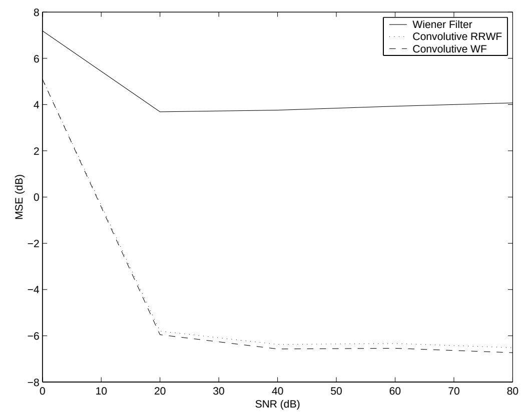

Figure 2: Graph comparing the performance of the Wiener filter, the convolutive (full rank) Wiener filter and the convolutive rank reduced Wiener filter when the time series $\boldsymbol{x}(t)$ was corrupted by noise.

8. REFERENCES

[1] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons, Inc., second edition, 1984.

[2] D. R. Brillinger. *Time Series: Data Analysis and Theory*. Holt, Rinehart and Winston, Inc., 1975.

[3] Y. Hua and M. Nikpour. Computing the reduced rank Wiener filter by IQMD. *IEEE Signal Processing Letters*, 6(9):240–242, September 1999.

[4] Y. Hua, M. Nikpour, and P. Stoica. Optimal reduced rank estimation and filtering. *IEEE Transactions on Signal Processing*, 2000. To appear.

[5] J. R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley and Sons, Inc., 1994.

[6] J. H. Manton. Optimisation algorithms exploiting complex orthogonality constraints. *IEEE Transactions on Signal Processing*, 2000. Submitted.

[7] J. H. Manton and Y. Hua. Rank reduction and James-Stein estimation. *IEEE Transactions on Signal Processing*, 47(11):3121–3125, November 1999.

[8] J. H. Manton, R. Mahony, and Y. Hua. The geometry of weighted low rank approximations. *IEEE Transactions on Signal Processing*, 2000. Submitted.

[9] L. L. Scharf. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley, 1991.

[10] L. L. Scharf. The SVD and reduced rank signal processing. *Signal Processing*, 25:113–133, 1991.

[11] L. L. Scharf and D. W. Tufts. Rank reduction for modeling stationary signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(3):350–355, March 1987.

[12] P. Stoica and M. Viberg. Maximum likelihood parameter and rank estimation in reduced-rank multivariate linear regressions. *IEEE Transactions on Signal Processing*, 44(12):3069–3078, December 1996.