On the relationship between whitened principal components analysis and maximum autocorrelation factors

Neal B. Gallagher, Jeremy M. Shaver, Randall Bishop, Robert T. Roginski, Barry M. Wise
SSC’13 17-20, June 2013
On the relationship between whitened principal components analysis and maximum autocorrelation factors

Maximum autocorrelation factors (MAF)\(^1\) and the use of de-cluttering algorithms have been gaining popularity recently in chemometrics. Generalized least squares weighting, developed by Aitken c. 1935\(^2\), is a popular and useful de-cluttering algorithm that can be used to "whiten" data prior to principal components analysis (PCA). The relationship between MAF and whitened PCA is elucidated with the objective of 'de-mystifying' the MAF algorithm. The insight is intended to provide a fundamental understanding of the methods such that they can be employed to best effect and that related methods can be easily developed using a similar relationship. Examples will be shown for hyperspectral images.

Compare PCA and WPCA

PCA = principal components analysis

WPCA = whitened PCA

Objective to Maximize

PCA \quad O(p) = p^T \Sigma p \quad | \quad p^T p = 1
\quad \Sigma = X^T X

WPCA \quad O(p) = p^T \Sigma_c^{-\frac{1}{2}} \Sigma \Sigma_c^{-\frac{1}{2}} p \quad | \quad p^T p = 1
\quad t = X \Sigma_c^{-\frac{1}{2}} p

\Sigma = X^T X
\Sigma_c = X_c^T X_c

Scores

\quad t = Xp

Clutter is interferences and noise; it is measured signal not of primary interest

X = data measurements
X_c = clutter measurements
Compare WPCA and MAF

WPCA = whitened PCA
MAF = maximum autocorrelation factors

Objective to Maximize

WPCA \( O(p) = p^T \Sigma_c^{-1/2} \Sigma \Sigma_c^{-1/2} p \ | \ p^T p = 1 \)

MAF \( O(w) = \frac{w^T \Sigma w}{w^T \Sigma_\Delta w} \)

Scores

\( t = X \Sigma_c^{-1/2} p \)

\( \Sigma = X^T X \)

\( \Sigma_\Delta = X^T D_1^T D_1 X \)

\( a = Xw \)
Define PCA, WPCA and MAF

PCA = principal components analysis
WPCA = whitened PCA
MAF = maximum autocorrelation factors

<table>
<thead>
<tr>
<th>Eigenvalue Problem</th>
<th>Eigenvalue ; Eigenvector</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA ( \Sigma \mathbf{p} = \lambda \mathbf{p} )</td>
<td>( \mathbf{p} ; \lambda )</td>
<td>( \mathbf{t} = \mathbf{Xp} )</td>
</tr>
<tr>
<td>WPCA ( \Sigma_c^{-1/2} \Sigma \Sigma_c^{-1/2} \mathbf{p} = \lambda \mathbf{p} )</td>
<td>( \mathbf{p} ; \lambda )</td>
<td>( \mathbf{t} = \mathbf{X} \Sigma_c^{-1/2} \mathbf{p} )</td>
</tr>
<tr>
<td>MAF ( \Sigma \mathbf{w} = \alpha \Sigma \Delta \mathbf{w} )</td>
<td>( \mathbf{w} ; \alpha )</td>
<td>( \mathbf{a} = \mathbf{Xw} )</td>
</tr>
</tbody>
</table>
Compare WPCA and MAF

**Eigenvalue Problem**

<table>
<thead>
<tr>
<th>Method</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>WPCA</td>
<td>$\Sigma_c^{-\frac{1}{2}} \Sigma \Sigma_c^{-\frac{1}{2}} p = \lambda p$</td>
</tr>
<tr>
<td>MAF</td>
<td>$\Sigma w = \alpha \Sigma \Delta w$</td>
</tr>
<tr>
<td>MSF</td>
<td>$\Sigma w = \alpha \Sigma_c w$</td>
</tr>
</tbody>
</table>

**Eigenvalue ; Eigenvector**

- $p ; \lambda$
- $w ; \alpha$
- $w ; \lambda$
- $w = \Sigma_c^{-\frac{1}{2}} p$

**Scores**

- $t = X \Sigma_c^{-\frac{1}{2}} p$
- $a = Xw$
- $t = Xw$
### MSF

<table>
<thead>
<tr>
<th>Eigenvalue Problem</th>
<th>Eigenvalue ; Eigenvector</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>WPCA</td>
<td>$\Sigma_c^{-\frac{1}{2}} \Sigma_s \Sigma_c^{-\frac{1}{2}} p = \lambda p$</td>
<td>$p ; \lambda$</td>
</tr>
<tr>
<td>MSF</td>
<td>$\Sigma_s w = \alpha \Sigma_c w$</td>
<td>$w ; \lambda$</td>
</tr>
<tr>
<td></td>
<td>$w = \Sigma_c^{-\frac{1}{2}} p$</td>
<td></td>
</tr>
</tbody>
</table>

$\Sigma_s = X_s^T D_s^T D_s X_s$

$\Sigma_c = X_c^T D_c^T D_c X_c$

Signal and clutter can include different data and spatial operators (or different row operators for time-series)
Six wires of different alloys in epoxy, measured with EDS.


Paul Kotula at Sandia and Mike Keenan are gratefully acknowledged for this data.
Partitioning of Variance

Variance captured by PC 1 is refocused

$$X_s ; D_s = I$$

$$X_c ; D_c = I$$
Target Detection

Variance is focused on Ni. The results are directly related to GLS target detection a.k.a. matched filter.

X_s; D_s = I

X_c; D_c = I

1 (top): 100% Ni
2: 36% Ni, 64% Fe

1Los Alamos National Laboratory, Los Alamos, NM; 2HORIBA Jobin Yvon, Inc, Edison, NJ; 3Eigenvector Research, Inc., Wenatchee, WA
MDF vs MAF

\[ X_s = X_c ; \quad D_s = D_1 ; \quad D_c = D_2 \]

\[ X_s = X_c ; \quad D_s = I ; \quad D_c = D_1 \]

Maximum Difference Factors is factor based edge detection
MDF

\[ X_s = X_c ; \quad D_s = D_1 ; \quad D_c = D_2 \]

movies can be used for visualization
Conclusions 1/2

- The generalized eigenvector problem (MAF, MSF, MDF…) can be reformed to a standard eigenvector problem (WPCA)
- Whitened PCA and MAF are mathematically similar for similar definition of signal and clutter
- PCA may be more interpretable (MAF “demystified”)

- The original formulation can be memory and time intensive. The GEP problem can be reformulated using subspace projections to save time and memory, and focus only on eigenvalues of interest.

\[
\Sigma_s = X_s^T D_s^T D_s X_s \quad \Sigma_c = X_c^T D_c^T D_c X_c
\]

\[
\hat{\Sigma}_s = T_s^T D_s^T D_s T_s \quad \hat{\Sigma}_c = T_c^T D_c^T D_c T_c
\]
Conclusions 2/2

- The spatial operators $D_i$ can be anything to focus the variance as desired
  - SavGol, wavelets, etc.
  - and can include variable operators

- Characterizing / measuring clutter is as important as characterizing the signal.
WPCA and MSF partition the variance $\Sigma$ through "re-weighting" to characterize clutter $\Sigma_c$ and signal of interest $\Sigma_s$.