Numerical errors in weight vector computation

Adam Bojanczyk
Electrical and Computer Engineering
Cornell University
Weight vector computation problem

Minimize array output in all directions but that of the target direction $d$

Minimization problem (1):

\[
\begin{align*}
\min \quad & w^H C w \\
\text{subject to} \quad & w^H d = 1
\end{align*}
\]

C is signal covariance matrix

Solution: \( w = \frac{C^{-1}d}{d^H C^{-1}d} \)
C is approximated by

\[ A = \mathbf{S}_k \mathbf{S}_k^H = \mathbf{X}^H \mathbf{X} \quad \text{where} \quad \mathbf{X} = [\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_K] \]

Minimization problem (2):

\[
\begin{align*}
\min & \quad \mathbf{w}^H \mathbf{A} \mathbf{w} \\
\text{subject to} & \quad \mathbf{w}^H \mathbf{d} = 1
\end{align*}
\]

A is sample covariance matrix

Solution: \( \mathbf{w} = \frac{\mathbf{A}^{-1} \mathbf{d}}{\mathbf{d}^H \mathbf{A}^{-1} \mathbf{d}} \)

\[ \mathbf{A} = \mathbf{X}^H \mathbf{X}. \quad \text{Minimization problem (3):} \]

\[
\begin{align*}
\min & \quad ||\mathbf{X} \mathbf{w}||_2 \\
\text{subject to} & \quad \mathbf{w}^H \mathbf{d} = 1
\end{align*}
\]

constrained least squares
Algorithms for (2)

\( A = X^H X > 0 \), need to compute \( A^{-1} d \)

<table>
<thead>
<tr>
<th>Normal Equations (NE)</th>
<th>Semi NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I) Cholesky</td>
<td>(II) GE</td>
</tr>
<tr>
<td>1) ( A = U^H U )</td>
<td>1) ( A = L^H U )</td>
</tr>
<tr>
<td>2) ( U^H u = d )</td>
<td>2) ( L^H u = d )</td>
</tr>
<tr>
<td>3) ( U v = u )</td>
<td>3) ( U v = u )</td>
</tr>
<tr>
<td>4) ( w = v / (d^H v) )</td>
<td>4) ( w = v / (d^H v) )</td>
</tr>
</tbody>
</table>

Computed \( A \) may become indefinite so use GE instead of Cholesky
Algorithms for (3) – Null Space Method

1) \( Q^H \begin{bmatrix} d^H \\ X \end{bmatrix} H = \begin{bmatrix} 0 & 0 \\ e_1^T & 0 \\ L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \)  
   Generalized QL decomposition

2) \( L_{22} w_2 = - L_{21} \)  
   \( L_{22} \) is better conditioned than \( X \)

3) \( w = H \begin{bmatrix} 1 & w_2^H \end{bmatrix}^H \)

4) \( ||Xw|| = |L_{11}| \)  
   \( L_{11} \) is the norm of the residual
Sensitivity Analysis - NE

A replaced by \( \hat{A} = A + ?A \), \( ||?A|| < ? ||A|| \)

\[
\begin{align*}
    w &= \frac{A^{-1}d}{d^H A^{-1}d} \\
    \hat{w} &= \frac{\hat{A}^{-1}d}{d^H \hat{A}^{-1}d}
\end{align*}
\]

Then

\[
\frac{||w - \hat{w}||}{||w||} < ? ||L_{22}^{-1}||^2 ||A||
\]
Sensitivity Analysis - SNE

$X$ replaced by $\hat{X} = X + \alpha X$, $\|\alpha X\| < \|X\|$

$$w = \frac{(X^H X)^{-1}d}{d^H (X^H X)^{-1}d} \quad \hat{w} = \frac{(\hat{X}^\top \hat{X})^{-1}d}{d^H (\hat{X}^\top \hat{X})^{-1}d}$$

$$\frac{\|w - \hat{w}\|}{\|w\|} < \alpha \left( \|L_{22}^{-1}\|^2 \|X\| \frac{\|Xw\|}{\|w\|} + \|L_{22}^{-1}\| \|X\| \right)$$

can be small
Sensitivity Analysis - NS

X replaced by \( \hat{X} = X + ?X \), \( ||?X|| < ?||X|| \)

\[
\min_{\hat{w}^{Hd} = 1} ||Xw||_2 \\
\min_{\hat{w}^{Hd} = 1} ||\hat{X} \hat{w}||_2
\]

\[
\frac{||w - \hat{w}||}{||w||} < ? \left( ||L_{22}^{-1}||^2 ||X|| \frac{||Xw||}{||w||} + ||L_{22}^{-1}|| ||X|| \right)
\]
SNE&NS - Small residual case

\[ X = W ? V^H, \ ? = \text{diag}(?i), \ V = [v_1, ..., v_n] \]

If \( d = v_n \) then \( w = v_n, \|Xw\| = ?_n \) and

\[
\frac{\|w - \hat{w}\|}{\|w\|} < ? \left( \|L_{22}^{-1}\|^2 \|X\| ?_n + \|L_{22}^{-1}\| \|X\| \right)
\]
Conclusions

• SNE and NS are equally accurate
  • this is not the case for general LS problems

• If $||Xw|| = ?_n$ then SNE and SN are more accurate than NE
  • in NE use GE instead of Cholesky

• If $||Xw|| = ?_1$ then NE, SNE and SN are equal
  • NE is the least expensive

• cond number determined by submatrices of L
  • cond number can be small even if that of X is large