Blind Source Separation Using Information Theoretic Learning

Kenneth Hild
April 12, 2000
Outline

• Blind Source Separation
• Information Theoretic Learning
  • Maximum Entropy (ME)
  • Minimum Mutual Information (MMI)
• Performance of MMI
• Comparison of MMI, ME and 2nd order methods
• Conclusions
Blind Source Separation

• Mixing \[ x = h^T s \]
• Demixing \[ y = w^T g^T x \]

- \( s \) - \((N \times L)\) source matrix
- \( h \) - \((N \times M)\) mixture matrix
- \( x \) - \((M \times L)\) observation matrix
- \( g \) - \((M \times M)\) whitening transform
- \( w \) - \((M \times N)\) rotation matrix
- \( y \) - \((N \times L)\) output matrix
Block Diagram

- Feedforward, instantaneous mixture model
- Gain/permutation ambiguities exist

Figure 1. Block diagram of overall system (including both mixing and demixing sub-blocks) when N=2 and M=2.
Maximum Entropy

- ME is unable to find the required rotation (w/o a nonlinearity)

\[ H(y) = -\int f_y(y) \ln f_y(y) dy \]

\[ f_y(y) = \frac{f_x(x)}{|\det w|} \]

\[ H(y) = H(x) + \ln |\det w| \]

\[ \max H(y) = \max |\det w| \]
Minimum Mutual Information

- MMI is able to find the required rotation

\[
I(y) = H(y_1) + H(y_2) - H(y)
\]

\[
\min I(y) = \min \log \frac{1}{N^2} \sum_i \sum_j G_1 + \log \frac{1}{N^2} \sum_i \sum_j G_2
\]

\[
G_1 = G(y_1^i - y_1^j, 2\sigma^2)
\]

\[
G_2 = G(y_2^i - y_2^j, 2\sigma^2)
\]
Minimum Mutual Information

- The tap weight update is a weighted sum of fields (weighted by the forces)

\[ \Delta \alpha = \left( \sum_i \sum_j G_2 \right) \left( \sum_i \sum_j \frac{\partial G_1}{\partial \alpha} \right) + \left( \sum_i \sum_j G_1 \right) \left( \sum_i \sum_j \frac{\partial G_2}{\partial \alpha} \right) \]

\[ w = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \]
Information Field/Force

- $N=100$, $\sigma=0.1$, optimum $\alpha=\pi/10$
Tap Weight Track

- $N=200$, $\sigma=0.3$, optimum $\alpha=\pi/4$
Recovered Sources

- MMI
- InfoMax
- 2nd order with nonlinearity
Conclusions

• Advantages of MMI
  • ITL is used to adapt 1 parameter instead of 4
  • Does not suffer from the curse of dimensionality
  • Does not require a nonlinearity
  • Able to find the rotation, unlike 2nd order methods

• Disadvantages of MMI
  • Greater computational complexity than 2nd order methods