Blind Deconvolution

and

Blind Source Separation

Shane M. Haas

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
I  – Introduction

II  – The Bussgang Algorithm for Blind De-convolution

III – An Information-Maximization Algorithm for Blind Source Separation

IV  – Demonstrations

V   – Conclusions
I – Introduction

- **THE PROBLEM:** Unraveling an unknown linear operation

Blind Deconvolution

![Diagram](attachment:diagram.png)

Blind Source Separation
The Channel Equalization Problem*

*Graphic from Institute for Communications Engineering Webpage
The Cocktail Party Problem*

*Graphic from Daniel Rowe's Webpage
• **A SOLUTION**: Blind adaptive signal processing
• Two Questions:

– How to choose the non-linearity?

– What should the adaptive mechanism optimize?
• **The Bussgang Algorithm for Blind De-convolution** (Haykin, 1991)
  
  – Non-Linearity: Bayes Estimator
  
  – Adaptive Mechanism: LMS Algorithm

• **Information-Maximization Method for Blind Source Separation** (Bell & Sejnowski, 1995)
  
  – Non-Linearity: Logistic Function
  
  – Adaptive Mechanism: InfoMax Algorithm
II – Bussgang Algorithm for Blind De-convolution

\[ y(n) = E[s(n)|u(n)] \]

\[ x(n) \rightarrow \text{Adjustable FIR Filter} \]

\[ \text{LMS Algorithm} \]

\[ u(n) \]

\[ y(n) \]

\[ w_k(n) \]
**The Non-Linearity:** Bayes Estimator

\[ u(n) = \sum_k w_k(n)x(n - k) \]
\[ = \sum_k a_k x(n - k) + \sum_k (w_k(n) - a_k) x(n - k) \]
\[ = s(n) + v(n) \]

where

\[ v(n) = \sum_k (w_k(n) - a_k) x(n - k) \]

\[ a_n = \text{Ideal Equalizing Filter} \]
• Assume that $v(n)$ is:

  – zero-mean,

  – Gaussian,

  – white, and is

  – statistically independent of the source sequence $s(n)$.

• Assume that $s(n)$ is iid uniform with zero mean and unit variance
• Under these assumptions:

\[ y(n) = E[s(n) | u(n)] \]
\[ = \frac{1}{c_0} u(n) + \frac{\sigma Z(u(n) + c_0/\sqrt{3}) - Z(u(n) - c_0/\sqrt{3})}{c_0 Q(u(n) - c_0/\sqrt{3}) - Q(u(n) + c_0/\sqrt{3})} \]

where \( \sigma^2 \) is the variance of \( v(n) \), \( c_0 = \sqrt{1 - \sigma^2} \) is a normalizing factor, and

\[ Z(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \]
\[ Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{z^2}{2}} dz \]
• **The Adaptive Mechanism:** The Least-Mean Squares (LMS) Algorithm

  - Adjust weights to minimize mean squared error:

    ![Diagram](image-url)
• Iterative approximation to optimal Wiener filter:

\[
\begin{align*}
\text{"Desired" Signal} & \quad y(n) \\
\text{Error Signal} & \quad e(n) \\
\text{Optimal Wiener Filter Projection} & \quad w_0^* x(n) + \ldots + w_{M-1}^* x(n-M+1) \\
\text{Subspace generated by} & \quad \{x(n), x(n-1), \ldots, x(n-M+1)\}
\end{align*}
\]
• Wiener-Hopf Equation:

\[ E[x(n)e^*(n)] = E[x(n)(y(n) - w^Hx(n))^*] \]
\[ = p - Rw \]
\[ = 0 \]

\[ \iff Rw = p \]

where

\[ w = (w_0, w_1, \ldots, w_{M-1})^T \]
\[ x(n) = (x(n), x(n-1), \ldots, x(n - M + 1))^T \]
\[ p = E[x(n)y^*(n)] \]
\[ R = E[x(n)x^H(n)] \]
• The Mean Squared Error

\[ J(n) = E \left[ |e(n)|^2 \right] \]

\[ = E \left[ |y(n) - u(n)|^2 \right] \]

\[ = E \left[ |y(n) - w^H(n)x(n)|^2 \right] \]

\[ = E[|y(n)|^2] - w^H(n)p \]

\[-p^Hw(n) + w^H(n)Rw(n)\]

• The Gradient (Assume \( y(n) \) is independent of \( w \))

\[ \frac{\partial J(n)}{\partial w(n)} = -2p + 2Rw(n) \]
• Steepest-Descent Algorithm

\[ \Delta w = w(n + 1) - w(n) \]

\[ \propto -\frac{\partial J(n)}{\partial w(n)} \]

\[ \propto p - Rw(n) \]

• The LMS Algorithm (p ≈ x(n)y*(n) and R ≈ x(n)x^H(n))

\[ \Delta w \propto x(n)y^*(n) - x(n)x^H(n)w(n) \]

\[ = x(n)(y(n) - w^H(n)x(n))^* \]

\[ = x(n)e^*(n) \]
III — An Information-Maximization Approach to Blind Source Separation

\[ x \rightarrow \text{Adjustable Linear Transformation} \rightarrow y = g(u) \]

\[ u = Wx + w \]
• The Non-Linearity: Logistic Function

\[ y_i = \frac{1}{1 + e^{-u_i}} \]

\[ u = Wx + w \]
• The Adaptive Mechanism: InfoMax Algorithm

− Maximize the differential entropy of the non-linear output

− The pdf of the non-linear output:

\[ f_Y(y) = \frac{f_X(x)}{|J|} \]

where \( J \) is the Jacobian

\[
J = \begin{vmatrix}
\frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_N} \\
\vdots & \ddots & \vdots \\
\frac{\partial y_N}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_N}
\end{vmatrix}
\]
– The differential entropy of the non-linear output:

\[
h(Y) = -E[\log f_Y(y)]
\]

\[
= E[\log |J|] - E[\log f_X(x)]
\]

– Adjust weights in recursive manner:

\[
\Delta W \propto \frac{\partial h(Y)}{\partial W} \approx \frac{\partial \log |J|}{\partial W}
\]

\[
\Delta w \propto \frac{\partial h(Y)}{\partial w} \approx \frac{\partial \log |J|}{\partial w}
\]
– Several pages of derivatives later...

\[ \Delta W \propto [W^T]^{-1} + (1 - 2y)x^T \]

\[ \Delta w \propto 1 - 2y \]
IV – Demonstrations
V – Conclusions

- Introduced blind deconvolution and blind source separation problem
- Bussgang algorithm for blind deconvolution
- Information-maximization algorithm for blind source separation
- Demonstrations