1 Problem Statement

Blind source separation (BSS) is a major area of research in signal and image processing [1]. It aims at recovering source signals from their mixtures without detailed knowledge of the mixing process.

The linear instantaneous mixing model of two source signals is

\[ X(t) = A_0 S(t) \]  \hspace{1cm} (1)

where \( S(t) \in R^2 \) is the time dependent source signal vector; \( A_0 \) is a 2x2 time independent unknown mixing matrix, \( X(t) \) is the known mixture signal. We wish to recover the source signal vector \( S \), without knowing \( A_0 \) (therefore blind), provided that all components of \( S(t) \) are independent random processes. In applications to sounds or images, the random processes in \( S(t) \) are non-Gaussian.

Solutions are non-unique in the sense that \( S(t) \) is undetermined up to scaling and permutation because of the product form \( A_0 S(t) \) of the unknowns. However, such degrees of freedom does not affect human perception of separation.

*Department of Mathematics, UCI, Irvine, CA 92617.*
2 Two by Two Instantaneous Mixtures

In component form, (1) is:

\[ x_1(t) = a_{11}s_1(t) + a_{12}s_2(t), \]
\[ x_2(t) = a_{21}s_1(t) + a_{22}s_2(t). \] (2)

To diagonalize (2), let us define:

\[ \nu_1(t) = a_{22}x_1(t) - a_{12}x_2(t), \]
\[ \nu_2(t) = -a_{21}x_1(t) + a_{11}x_2(t), \] (3)

then:

\[ \nu_1(t) = (a_{11}a_{22} - a_{12}a_{21})s_1(t) = \det(A_0)s_1(t), \]
\[ \nu_2(t) = \det(A_0)s_2(t). \] (4)

Assume that \( A_0 \) is non-singular \( (\det(A_0) \neq 0) \), we see that processes \( \nu_1(t) \) and \( \nu_2(t) \) are independent of each other. Now we approximate the independence condition by cross correlations.

2.1 Decorrelation with Second Order Statistics

Suppose signals are stationary over a couple of time frames of 5 to 10 ms length, we have:

\[ E[\nu_1(t) \nu_2(t - n)] = 0, \quad n \in [-N, N], \] (5)

where the expectation is approximated by sample average using data in the frames.

Substituting (3) in (5), we obtain:

\[ 0 = E[\nu_1(t) \nu_2(t - n)] \]
\[ = E[(a_{22}x_1(t) - a_{12}x_2(t))(-a_{21}x_1(t - n) + a_{11}x_2(t - n))] \]
\[ = -a_{22}a_{21}C_{11}^{n} + a_{12}a_{21}C_{21}^{n} + a_{22}a_{11}C_{12}^{n} - a_{12}a_{11}C_{22}^{n}. \] (6)
where $C_{n}^{ij} = E[x_i(t)x_j(t-n)]$ are known from received data. Normalizing amplitudes and introducing angle variables:

$$a^{22} = \cos(\theta), \ a^{12} = \sin(\theta),$$
$$a^{21} = \cos(\phi), \ a^{11} = \sin(\phi),$$ \hspace{1cm} (7)

we write equation (6) as:

$$C_{n}^{11} - \tan(\theta) C_{n}^{21} - \tan(\phi) C_{n}^{12} + \tan(\theta) \tan(\phi) C_{n}^{22} = 0.$$ \hspace{1cm} (8)

Letting $n = 1, 2$, we have two equations for two knowns:

$$C_1^{11} - \tan(\theta) C_1^{21} - \tan(\phi)[C_1^{12} - \tan(\theta)C_1^{22}] = 0,$$
$$C_2^{11} - \tan(\theta)C_2^{21} - \tan(\phi)[C_2^{12} - \tan(\theta)C_2^{22}] = 0.$$ \hspace{1cm} (9)

implying:

$$\tan(\phi) = \frac{C_1^{11} - \tan(\theta)C_1^{21}}{C_1^{12} - \tan(\theta)C_1^{22}} = \frac{C_2^{11} - \tan(\theta)C_2^{21}}{C_2^{12} - \tan(\theta)C_2^{22}},$$ \hspace{1cm} (10)

or:

$$a \tan^2(\theta) + b \tan(\theta) + c = 0,$$ \hspace{1cm} (11)

where:

$$a = C_1^{21}C_2^{22} - C_1^{22}C_2^{21},$$
$$b = C_1^{22}C_2^{11} + C_1^{12}C_2^{21} - C_1^{21}C_2^{12} - C_1^{11}C_2^{22},$$
$$c = C_1^{11}C_2^{12} - C_1^{12}C_2^{11}.$$

If $A_0$ is not nearly singular, equation (11) in general has two real solutions:

$$\tan(\theta) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$
and corresponding $\tan(\phi)$ follows from (10).

If $b^2 < 4ac$, one can sample $n$ differently, e.g. pick $n = 0, 2$ or $n = -1, 1$ and search for real solutions nearby. The other way to go is to solve (9) approximately by minimizing the sum of squares of the left hand side. Because (9) is quadratic, the cost function of such least square minimization problem is a polynomial of degree 4, its derivative is a cubic polynomial so the critical point equation has a real solution in closed analytical form. The minimization problem is solvable by an elementary algebraic formula. One may also call optimization routines in Matlab to find numerical solutions.

A partial solution is to minimize (11). Suppose $a > 0$, $b^2 < 4ac$, then the left hand side is minimized at $\tan(\theta) = -b/(2a)$, and the corresponding $\tan(\phi)$ is given by one of the equations of (9). Solution is the same if $a < 0$. One equation in (9) is satisfied exactly, the other approximately.

We see that solving BSS is about finding angles (directions), similar to eigenvalue problems in linear algebra!

The above method is called “decorrelation”. It uses second order statistics with the help of time shift to generate enough equations.

### 2.2 Demixing with joint 2nd and 3rd Order Statistics

The two equations are:

$$E[\nu_1(t) \nu_2(t)] = 0, \ E[\nu_1^2(t) \nu_2(t)] = 0. \quad (12)$$

Upon substitution by (3), we find (6) and:

$$-(a^{22})^2 a^{21} C^{111} + 2a^{22} a^{12} a^{21} C^{121} - (a^{12})^2 a^{21} C^{221} + (a^{22})^2 a^{11} C^{112} - 2a^{22} a^{12} a^{11} C^{122} + (a^{12})^2 a^{11} C^{222} = 0. \quad (13)$$
In angle variables (7), equation (13) is:
\[
C^{111} - 2 \tan(\theta) C^{121} + \tan^2(\theta) C^{221} - \tan(\phi)(C^{112} - 2 \tan(\theta) C^{122} + \tan^2(\theta) C^{222}) = 0. \tag{14}
\]
Combining (8) and (14), we get:
\[
\frac{C^{11} - \tan(\theta) C^{21}}{C^{12} - \tan(\theta) C^{22}} \tan(\phi) = \frac{C^{111} - 2 \tan(\theta) C^{121} + \tan^2(\theta) C^{221}}{C^{112} - 2 \tan(\theta) C^{122} + \tan^2(\theta) C^{222}} \tag{15}
\]
implying \((X = \tan(\theta))\):
\[
aX^3 + bX^2 + cX + d = a(X - X_1)(X - X_2)(X - X_3) = 0, \tag{16}
\]
where:
\[
a = C^{22}C^{221} - C^{21}C^{222}, \\
b = C^{11}C^{222} + 2C^{21}C^{122} - C^{12}C^{221} - 2C^{22}C^{121}, \\
c = 2C^{12}C^{121} + C^{22}C^{111} - 2C^{11}C^{122} - C^{21}C^{112}, \\
d = C^{11}C^{112} - C^{12}C^{111};
\]
and the three roots are:
\[
X_1 = S + T - \frac{b}{3a} \\
X_2 = -\frac{S + T}{2} - \frac{b}{3a} + \frac{\sqrt{3}}{2}(S + T)i, \\
X_3 = -\frac{S + T}{2} - \frac{b}{3a} - \frac{\sqrt{3}}{2}(S + T)i,
\]
where:
\[
S = (R + \sqrt{Q^3 + R^2})^{1/3}, T = (R - \sqrt{Q^3 + R^2})^{1/3}, \\
Q = \frac{3ac - b^2}{9a^2}, R = \frac{9abc - 27a^2d - 2b^3}{54a^3}. \tag{17}
\]
If all three roots are real, pick the one to minimize $|E[\nu_1(t)\nu_2^2(t)]|$. Once $\tan(\theta)$ is determined, (15) yields $\tan(\phi)$, and finally $a^{ij} (i, j = 1, 2)$.

To assess the separation quality of an algorithm, we define an objective measure called signal-to-interference-ratio improvement (SIRI) by measuring the crosstalk reduction before ($SIR_i$) and after ($SIR_o$) the demixing. Let the demixing matrix be:

$$W = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} = \begin{bmatrix} \tilde{a}^{22} & -\tilde{a}^{12} \\ -\tilde{a}^{21} & \tilde{a}^{11} \end{bmatrix}$$

where tildes denote the estimated values. In the mixing model (2), each channel contains direct and cross talk components:

$$(\text{direct}) \ x_{11}(t) = a_{11}^1 s_1(t), \quad (\text{cross}) \ x_{12}(t) = a_{12}^1 s_2(t),$$
$$(\text{cross}) \ x_{21}(t) = a_{21}^2 s_1(t), \quad (\text{direct}) \ x_{22}(t) = a_{22}^2 s_2(t). \quad (18)$$

The demixed output is:

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} w_{11}a_{11} + w_{12}a_{21} & w_{11}a_{12} + w_{12}a_{22} \\ w_{21}a_{11} + w_{22}a_{21} & w_{21}a_{12} + w_{22}a_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

The direct and cross talk components in the output are:

$$(\text{direct}) \ u_{11}(t) = (w_{11}a_{11} + w_{12}a_{21})s_1(t),$$
$$(\text{cross}) \ u_{12}(t) = (w_{11}a_{12} + w_{12}a_{22})s_2(t),$$
$$(\text{cross}) \ u_{21}(t) = (w_{21}a_{11} + w_{22}a_{21})s_1(t),$$
$$(\text{direct}) \ u_{22}(t) = (w_{21}a_{12} + w_{22}a_{22})s_2(t). \quad (19)$$

The input SIR in decibel (dB) is:

$$SIR_i = 10 \log_{10} \left[ \frac{\|x_{11}(\cdot)\|_2}{\|x_{12}(\cdot)\|_2} + \frac{\|x_{22}(\cdot)\|_2}{\|x_{21}(\cdot)\|_2} \right]$$
the output SIR in dB is:

\[
SIR_0 = 10 \log_{10} \left[ \frac{\|u_{11}(\cdot)\|_2}{\|u_{12}(\cdot)\|_2} + \frac{\|u_{22}(\cdot)\|_2}{\|u_{21}(\cdot)\|_2} \right]
\]

where \( \| \cdot \|_2 \) denotes the vector \( l_2 \) norm. The SIRI in dB is:

\[
SIRI = SIR_0 - SIR_i. \tag{20}
\]

References