



THE FAST ICA ALGORITHM

(OVERVIEW)

R. Frank K. Glass

Neuroinformatics Center



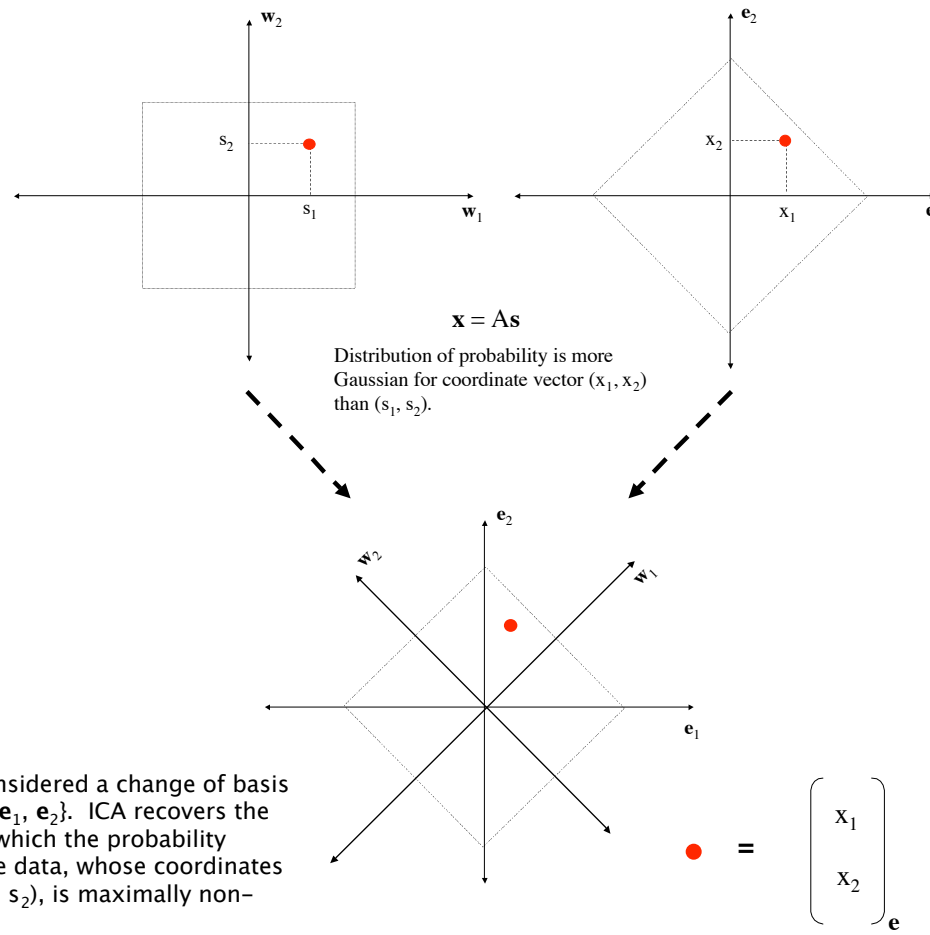
What is ICA?

ICA separates n statistically independent inputs, with non-Gaussian distributions, which have been linearly mixed into n output channels.

Statement of the Problem

- \mathbf{s} = the random vector of statistically independent components s_1, \dots, s_n .
- \mathbf{x} = the random vector whose elements are the mixtures x_1, \dots, x_n .
- A is the unknown 'mixing' matrix.
- $x_i = a_{i1}s_1 + \dots + a_{in}s_n$, for $i = 1 \dots n$,
or
 $\mathbf{x} = A\mathbf{s}$.
- ICA 'finds' the inverse of A , $W = A^{-1}$, such that
 $\mathbf{s} = W\mathbf{x}$.

Theoretical Basis, Heuristically Explained...



$\mathbf{x} = A\mathbf{s}$ can be considered a change of basis from $\{\mathbf{w}_1, \mathbf{w}_2\}$ to $\{\mathbf{e}_1, \mathbf{e}_2\}$. ICA recovers the basis $\{\mathbf{w}_1, \mathbf{w}_2\}$ in which the probability distribution of the data, whose coordinates are the vector (s_1, s_2) , is maximally non-Gaussian.

Conclusion

A result of the Central Limit Theorem is that the probability distribution of a linear combination of statistically independent random variables is more 'Gaussian' than any of the random variables which comprise the sum. Since $x_i = a_{i1}s_1 + \dots + a_{in}s_n$, $\text{pdf}(x_i)$ is more Gaussian than $\text{pdf}(s_i)$, $i = 1 \dots n$.

If each row, \mathbf{w}_i^T , of the unmixing matrix W is computed to minimize the non-Gaussianity of $\mathbf{w}_i^T\mathbf{x}$, for $i = 1 \dots n$, then we will have extracted the independent components. These components constitute the $\{\mathbf{w}\}$ basis.

$$\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} = \mathbf{s} = W\mathbf{x} = \begin{pmatrix} \mathbf{w}_1^T\mathbf{x} \\ \vdots \\ \mathbf{w}_n^T\mathbf{x} \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_e = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}_w$$