BMT 922 : Neural Signal Analysis and Modeling
Slides Series 2
PCA & ICA

Prof. Dr. Dr. Daniel J. Strauss

Eigenvalues and eigenvectors

For this entire Lecture, let A be square ($m \times m$). Let $x \neq 0 \in \mathbb{R}^m$.

Then x is an *eigenvector* of A and $\lambda \in \mathbb{R}$ is its corresponding *eigenvalue* if

$$Ax = \lambda x$$

The idea is that the action of A on a subspace S of $\mathbb{R}^m$ can act like scalar multiplication.

This special subspace S is called an *eigenspace*.

The set of all the eigenvalues of a matrix A is called the *spectrum* of A, denoted $\Lambda(A)$.

Characteristic polynomial

The *characteristic polynomial* p_A of A is the degree-m polynomial

$$p_A(z) = \det(zI - A).$$

Theorem
λ is an eigenvalue of A if and only if $p_A(\lambda) = 0$.

Note
Even in A is real, λ could be complex!
However, if A is real, any complex λ must appear in complex conjugate pairs.
i.e., if A is real and $\lambda = a + ib$ is an eigenvalue, then so is $\lambda^* = a - ib$.

Principal Compnent Analysis (PCA)

Principal Component Analysis (PCA)

(1) reduction of the number of variables in a data set (without significant loss in information)

(2) identification of new variables with greater meaning

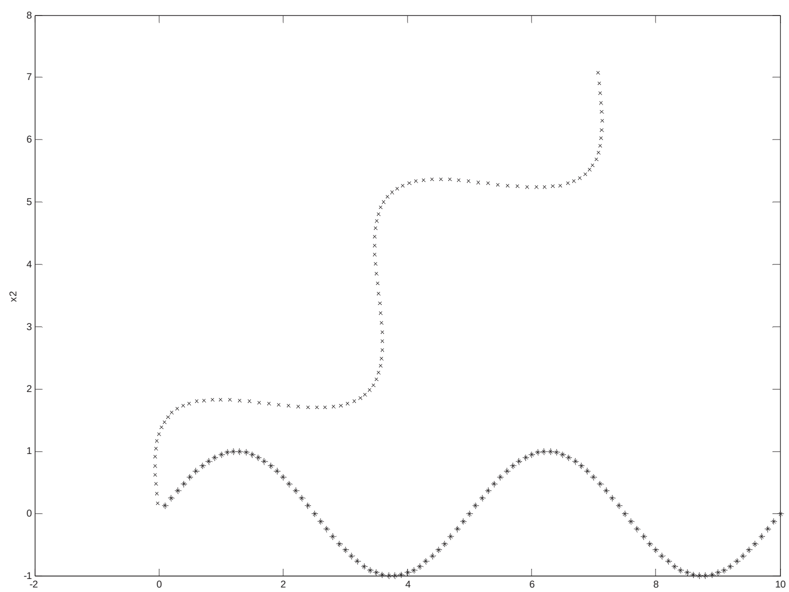
(3) transforms correlated variables in uncorrelated variables

(4) no value if the data is already uncorrelated

(5) uncorrelated vs. independent

Systems Neuroscience & Neurotechnology Unit

Data Rotation



rotation of data by 45°

Independent Compnent Analysis (ICA)

Systems Neuroscience & Neurotechnology Unit

Comparison PCA and ICA

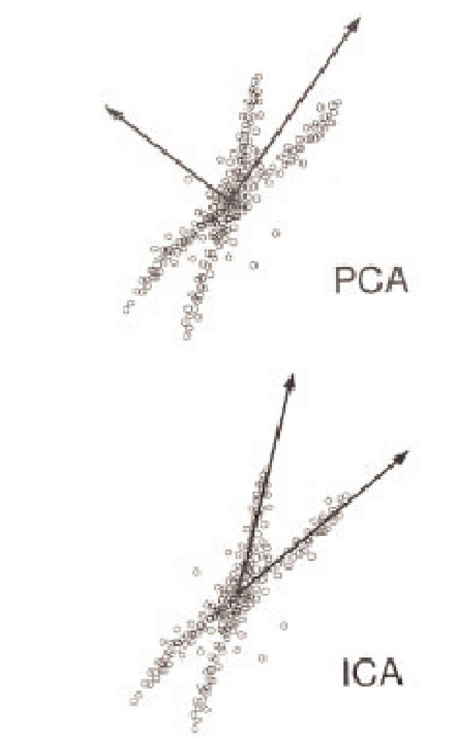


Fig. 1. The difference between PCA and ICA on a nonorthogonal mixture of two distributions that are independent and highly sparse (peaked with long tails). An example of a sparse distribution is the Laplacian: $p(x) = ke^{-|x|}$. PCA, looking for orthogonal axes ranked in terms of maximum variance, completely misses the structure of the data. Although these distributions may look strange, they are quite common in natural data.

assumptions of the original ICA

(1) sources are statistically independent

(2) at most one source with Gaussian distribution

(3) signals at the sensors are different linear combinations of the sources

(4) no time delay from sources to sensors

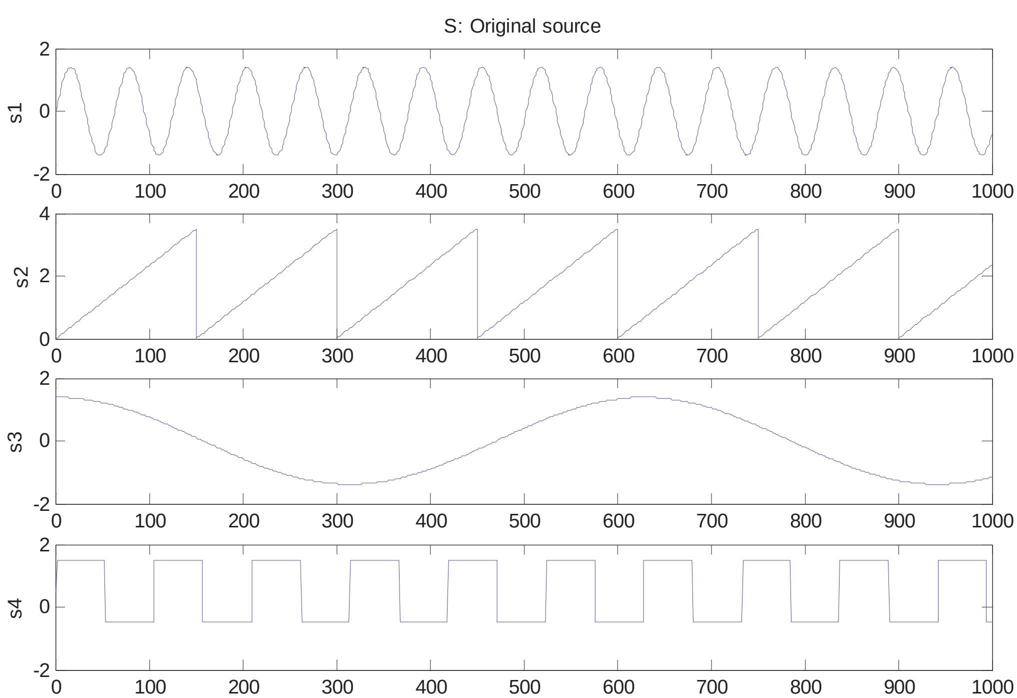
(5) same number of sources and sensors

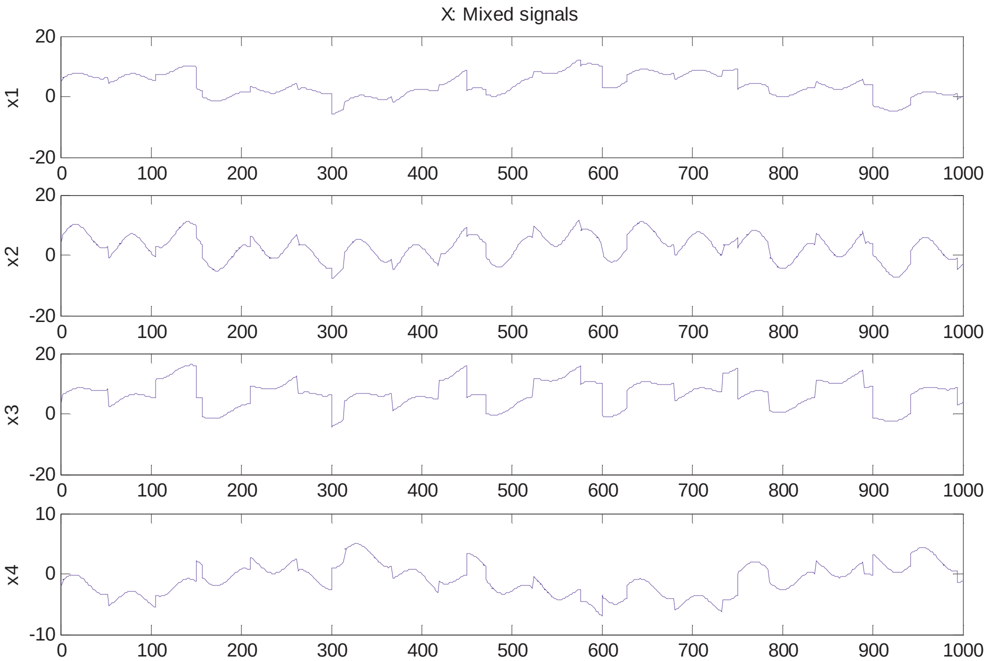
often not satisfied in real world applications

-

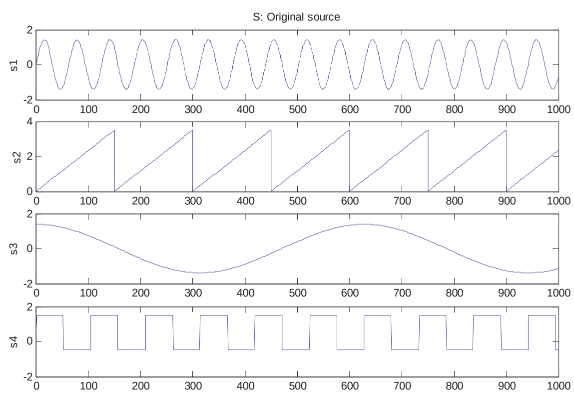
nevertheless, reasonable estimates can be obtained with the ICA

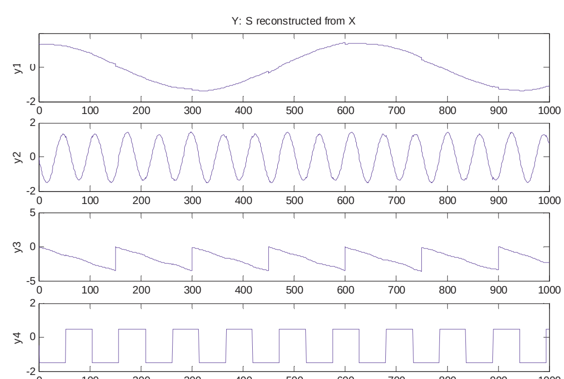
Systems Neuroscience & Neurotechnology Unit



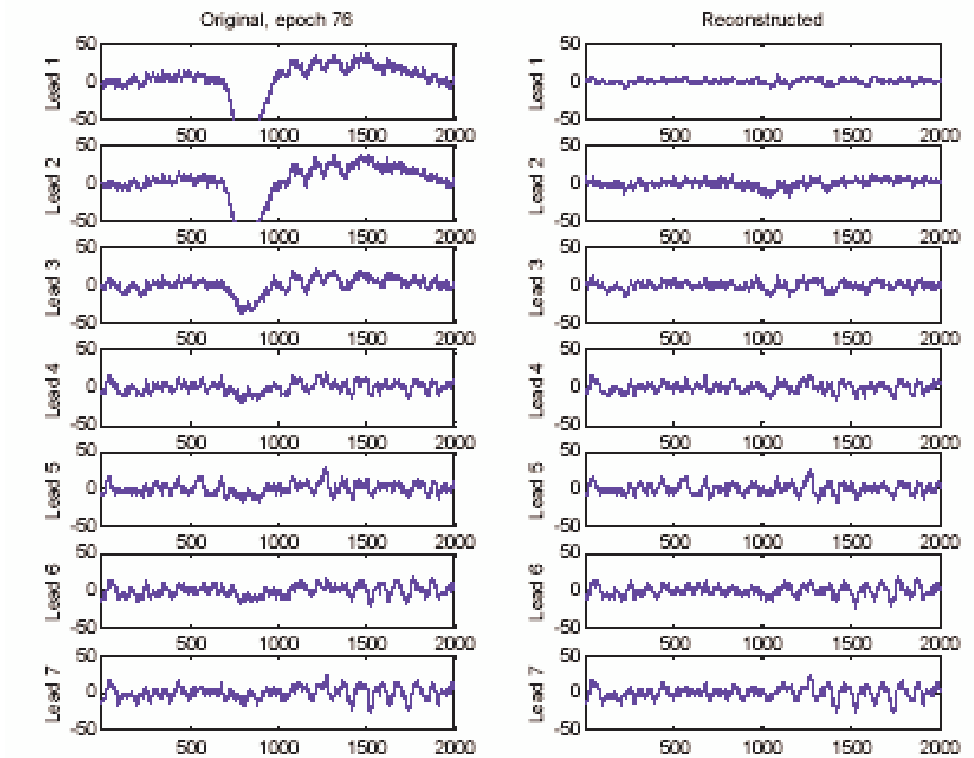


Systems Neuroscience & Neurotechnology Unit

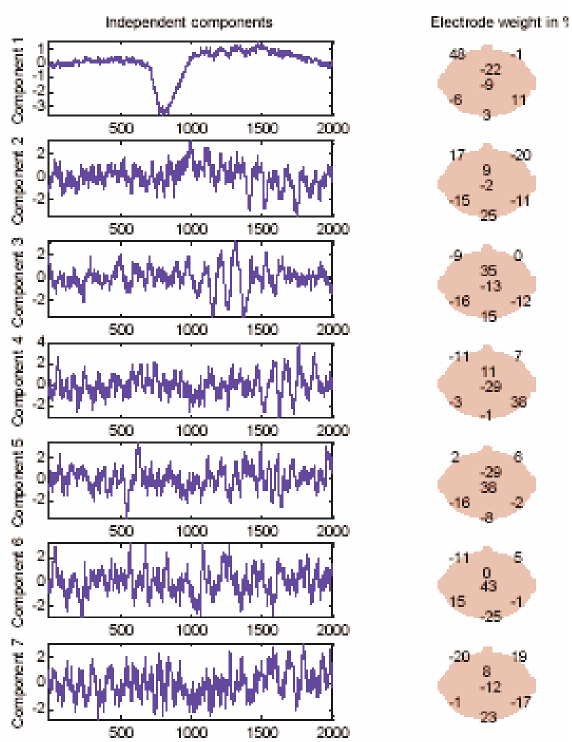




artifact extraction



Systems Neuroscience & Neurotechnology Unit



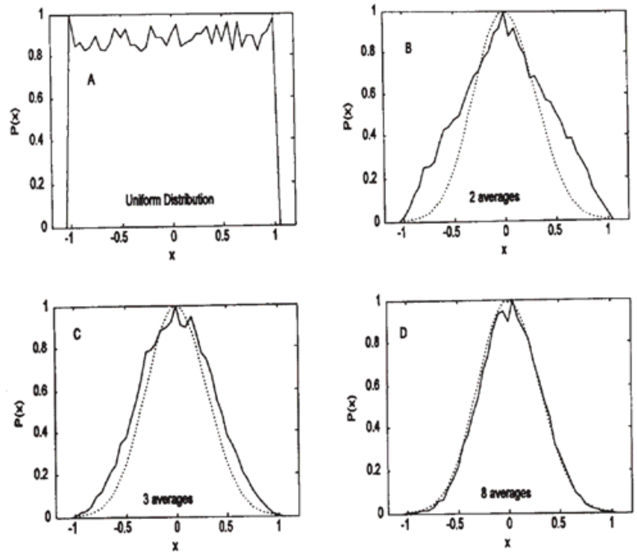


Figure 2.1 (A) The distribution of 20,000 uniformly distributed random numbers. (B) The distribution of 20,000 numbers, each of which is the average of two uniformly distributed random numbers. (C) and (D) The distribution obtained when 3 and 8 random numbers, still uniformly distributed, are averaged together. Although the underlying distribution is uniform, the averages of these uniformly distributed numbers tend toward a Gaussian distribution (dotted line). This is an example of the Central Limit Theorem at work.