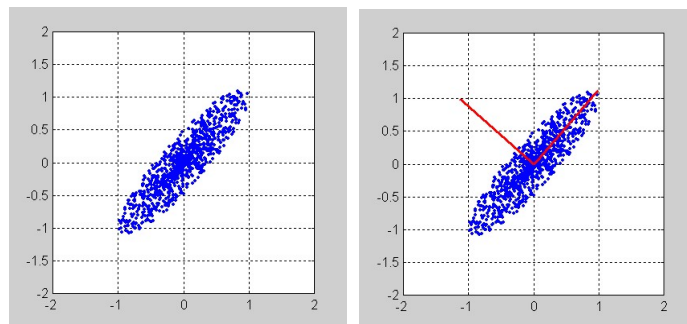


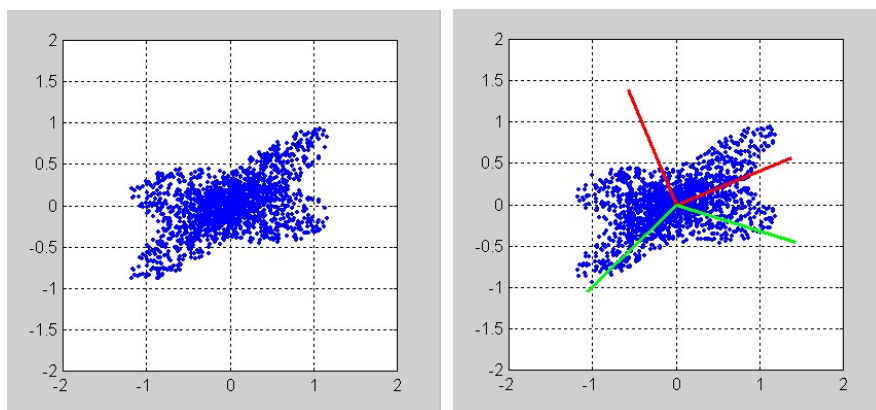
## (2) Independent Component Analysis (DHS 10.13.3)

- While PCA seeks directions of feature space that best represent the data in a sum-squared error sense, ICA seeks directions that are most statistically independent from each other.
- ICA seeks for solution to Blind Source Separation.
- What is Blind Source Separation?
  - Suppose  $k$ -dimensional data is observed at each moment  
 $s(t) = Ax(t)$   
where  $A$  is  $k \times d$  matrix,  $x$  sources, and  $s$  measured signals
  - Goal is to find  $W$  such that  $x(t) \approx Ws(t)$ , but  $A$  and  $x(t)$  are unknown in advance.

### PCA



### ICA



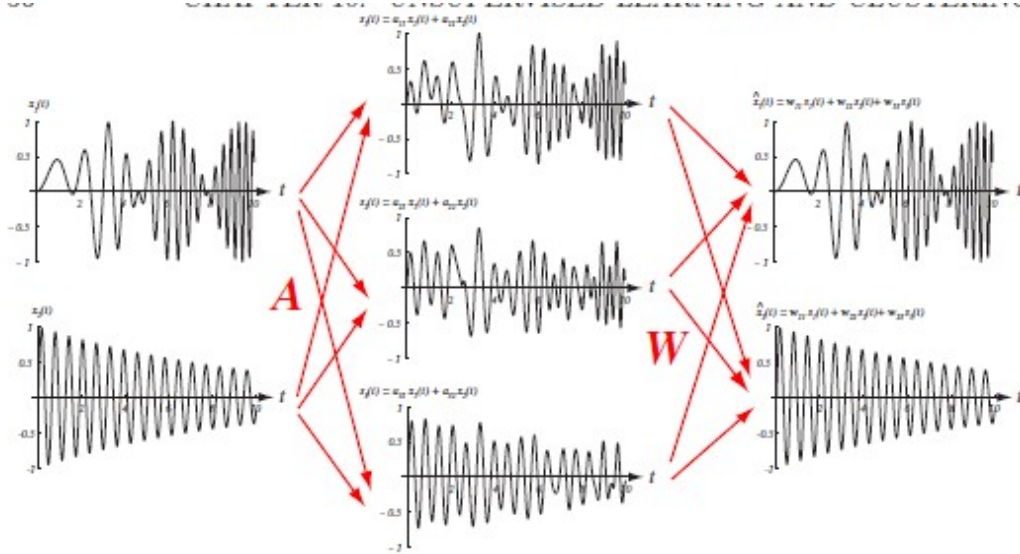
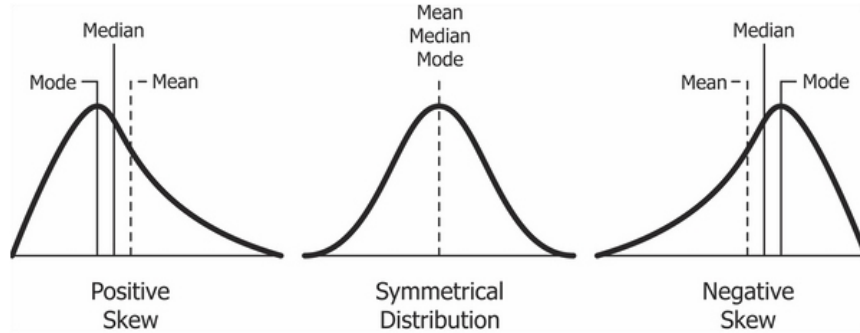


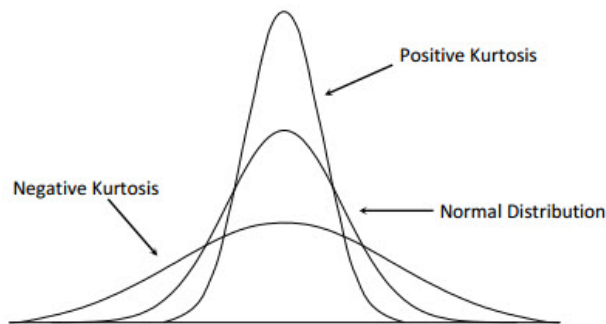
Figure 10.24: Independent component analysis (ICA) is an unsupervised method that can be applied to the problem of blind source separation. In such problems, two or more source signals (assumed independent)  $x_1(t)$ ,  $x_2(t)$ ,  $\dots$ ,  $x_d(t)$  are combined to yield a sum signal,  $s_1(t) + s_2(t) + \dots + s_c(t)$  where  $c \geq d$ . (This figure illustrates a case with only two components.) Given merely the linear signals, and the assumption of the number of components,  $d$ , the task of ICA is to recover the source signals. This is equivalent to finding a matrix  $\mathbf{W}$  that is the inverse of  $\mathbf{A}$ . In general applications of ICA, one seeks to extraction independent components from the sensed signals, whether or not they arose from a linear mixture of initial sources.

- ICA solutions require understandings up to the 4<sup>th</sup>-order statistics.

- First moment: mean  $m_x$
- Second moment: variance  $\sigma_x^2$
- Third moment:  $E\{x^3\}$  called skewness, a measure of asymmetry of the pdf.



- Fourth moment:  $E\{x^4\}$  called kurtosis, a measure of peakedness of the pdf.



- Also understanding on entropies, negentropy, and mutual information from Information Theory (DHS A.7). They are beyond the scope of this class.
- History
  - Works started in the early 1980s
  - Seminal work in the mid-90s by A.J. Bell and T.J. Sejnowski
  - Explosion of the research works since 1990s.
  - Now international workshops and more than hundreds researches work on the field.
  - Most popular algorithm: Fast ICA