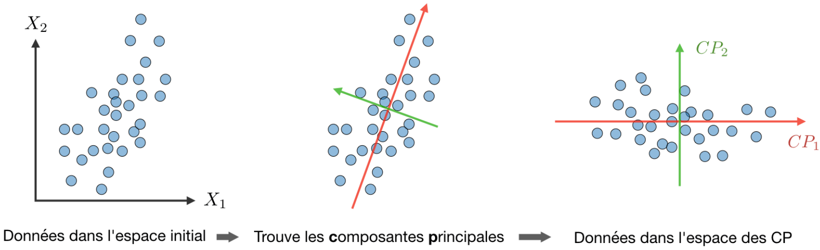
— Étape 4 : Projeter les données sur $\text{span}_{\mathbb{R}}(u_1,...,u_k)$. Cette procédure maximise la variance sur tous les espaces à k dimensions.



2.3.2 Analyse en composantes indépendantes

C'est une technique qui vise à trouver les sources génératrices sous-jacentes.

❒ **Hypothèses** – On suppose que nos données x ont été générées par un vecteur source à n dimensions $s = (s_1,...,s_n)$, où les s_i sont des variables aléatoires indépendantes, par le biais d'une matrice de mélange et inversible A de la manière suivante :

$$\boxed{x = As}$$

Le but est de trouver la matrice de démélange $W = A^{-1}$.

❒ **Algorithme d'ICA de Bell and Sejnowski** – Cet algorithme trouve la matrice de démélange W en suivant les étapes ci-dessous :

— Écrire la probabilité de $x = As = W^{-1}s$ par :

$$p(x) = \prod_{i=1}^{n} p_s(w_i^T x) \cdot |W|$$

— Écrire la log vraisemblance de notre jeu de données d'entrainement $\{x^{(i)}, i \in [\![1,m]\!]\}$ et en notant g la fonction sigmoïde par :

$$l(W) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \log \Big(g'(w_j^T x^{(i)}) \Big) + \log |W| \right)$$

Par conséquent, l'algorithme du gradient stochastique est tel que pour chaque example du jeu d'entrainement $x^{(i)}$, on met à jour W de la manière suivante :

$$\boxed{W \longleftarrow W + \alpha \left(\begin{pmatrix} 1 - 2g(w_1^T x^{(i)}) \\ 1 - 2g(w_2^T x^{(i)}) \\ \vdots \\ 1 - 2g(w_n^T x^{(i)}) \end{pmatrix} {x^{(i)}}^T + (W^T)^{-1} \right)}$$