5 **(PCA and ICA)** Let the n-dimensional random vector X be generated by a linear mixing model $X = AS$, with A be a $n \times n$ invertible matrix, $S = (S_1, \dots, S_n)^T$ with the components S_i mutually independent, centered, and variance 1. Let $\tilde{X} = VX$ be a whiten version of X ($E[\tilde{X}] = 0$ and $Cov[\tilde{X}] = I$). Write $\tilde{X} = RS$ with R an orthogonal matrix.

(a) **Non-identifiability with Gaussian sources (connection to PCA)**
Assume all S_i are Gaussian. Show that for any orthogonal Q, the transformed output $Y = Q^T\tilde{X}$ has independent components. Conclude that ICA is not identifiable (there is no unique unmixing) in this case, and explain why PCA is the best we can do (up to rotations in eigenspaces with equal eigenvalues) in the sense that no method, ICA included, can identify a unique unmixing beyond the covariance structure.

(b) **Identifiability with non-Gaussianity**
Assume at most one of the S_i is Gaussian, the rest are non-Gaussian. Show that if an orthogonal matrix W yields $Y = W\tilde{X}$ with independent components, then W must equal a signed permutation of R^T.
(Hint: The Darmois–Skitovich Theorem can be used without proof.)

(a) $\tilde{X} = RS$, $S \sim N(0, I)$, R orthogonal $\Rightarrow RR^T = I$

For any orthogonal Q, $Q^TQ = I$, $Y = Q^T\tilde{X} = Q^TRS = (Q^TR)S$, $S \sim N(0,I)$ so Y still normal

$Cov(Y) = (Q^TR)Cov(S)(Q^TR)^T = (Q^TR)I(Q^TR)^T = Q^T(RR^T)Q = Q^TQ = I \Rightarrow Y \sim N(0,I)$

$\Rightarrow$ Components are independent for $Cov(Y)$ is diagonal (in Normal, uncorrelate = independent)

In ICA, both S and any rotated version Y are "component independent", S and Y cannot be extinguished

In PCA, only 2nd order statistics covariance are used, any higher order statistics are not. In Gaussian, only mean and covariance are useful, no higher order information are contained, so PCA capture all information in Gaussian

(b) $Y = W\tilde{X}$, $\tilde{X} = RS \Rightarrow Y = WRS$, let $U = WR$, then $UU^T = I \Rightarrow Y_j = \sum_{i=1}^{n} U_{ji}S_i$

Darmois–Skitovich: $X_1, \dots, X_n$ mutually independent, $Y_1 = \sum_{j=1}^{n} a_jX_j$, $Y_2 = \sum_{j=1}^{n} b_jX_j$, if $Y_1 \perp Y_2$, X_j non-Gaussian, then $a_jb_j = 0$

Pick 2 different component Y_j, Y_k, $Y_j = \sum_{i=1}^{n} U_{ji}S_i$ $Y_k = \sum_{i=1}^{n} U_{ki}S_i \Rightarrow$ If S_i non-Gaussian, $U_{ji}U_{ki} = 0$

Only 1 S_i is Gaussian $\Rightarrow$ at least $n-1$ S_i follow that

$U_{jk}U_{ki} = 0 \Rightarrow$ For non-Gaussian column i, any two rows j, k cannot both non-zero $\Rightarrow$ only 1 non-zero (at most)

U orthogonal, every column's norm is 1, so in non-Gaussian column, there's only a 1 or -1 in the column (other are 0)

To be orthogonal to other $n-1$ columns, the Gaussian column also has only one 1 or -1 (other are 0)

$\Rightarrow U$ is a signed permutation matrix $\Rightarrow U = WR \Rightarrow W = UR^T$

$\Rightarrow W$ is a signed permutation of R^T