6 **(Factor analysis and PCA)** Generate $n = 100$ observations of the 20-dim vector Y from the following factor analysis model:

$$Y = \mu^* + \Lambda^* X + W$$

where X and W are independent, $\mu^* = \mathbf{0}_{20\times 1}$ and

$$\Lambda^* = \begin{pmatrix} \mathbf{1}_{10\times1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{10\times1} \end{pmatrix}_{20\times 2}, X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}_{2\times1} \sim N_2(0, I_2), \quad W_{20\times1} \sim N_{20}(0, 0.5\times I_{20}).$$

(a) Show that Λ can only be identified up to an orthornormal transformation. Propose one suitable distance measure to evaluate the distance of an estimated *column space* of Λ to the true column space (of Λ^*).

(b) Suppose we know that there are two factors in the factor analysis model. Fit the factor analysis model, and what is the equation that defines the estimated subspace of scores X? Plot the projections of the data set onto the 2-dim subspace.

(c) Implement PCA to identify the principal components and the projections of the data set on to the 2-dim principal subspace. Compare your results with those obtained in part (b).

(a) $\Sigma = Cov(Y) = Cov(\Lambda X + W) = \Lambda\, Cov(X)\Lambda^T + Cov(W)$

$Cov(X) = I,\ X \perp W \Rightarrow \Sigma = \Lambda\Lambda^T + Cov(W)$

Let Q orthogonal, $Q^TQ = I$, $\Rightarrow \tilde{X} = Q^TX$, $\tilde{\Lambda} = \Lambda Q \Rightarrow Y = \tilde{\Lambda}\tilde{X} + W = (\Lambda Q)(Q^TX) + W = \Lambda X + W$

$\Rightarrow Y$ doesn't change

$E[\tilde{X}] = E[Q^TX] = Q^TE[X] = 0$, $Cov(\tilde{X}) = Q^TCov(X)Q = Q^TIQ = I \Rightarrow \tilde{X} \sim N(0, I)$

$\Rightarrow \tilde{\Sigma} = \tilde{\Lambda}\tilde{\Lambda}^T + Cov(W) = (\Lambda Q)(\Lambda Q)^T + Cov(W) = \Lambda(QQ^T)\Lambda^T + Cov(W) = \Lambda\Lambda^T + Cov(W) = \Sigma$

$\Rightarrow \Sigma$ is same to any orthogonal transformation

Distance measure: Let P_{Λ^*} be the projection matrix to the true column space
$P_{\hat{\Lambda}}$ be the projection matrix to the estimated column space

$P_{\Lambda^*} = \Lambda^*(\Lambda^{*T}\Lambda^*)^{-1}\Lambda^{*T}$, $P_{\hat{\Lambda}} = \hat{\Lambda}(\hat{\Lambda}^T\hat{\Lambda})^{-1}\hat{\Lambda}^T$, $\text{Distance} = \|P_{\Lambda^*} - P_{\hat{\Lambda}}\|_F$

For projection matrix only decided by subspace instead of rotation