Implementation in MATLAB of the Partial Least Squares algorithm for classification

Case study: fault detection and diagnosis on steel plates

Lorenzo Ferrari Lorenzo Leoni

Department of Engineering and Applied Sciences, University of Bergamo

April 28, 2023

Outline

Introduction to the PLS technique

Partial least squares (**PLS**), as known as **projection to latent structures**, is a dimensionality reduction technique for maximizing the **covariance** between the predictor (independent) matrix $X \in \mathbb{R}^{n \times m}$ and the predicted (dependent) matrix $Y \in \mathbb{R}^{n \times p}$ for each component of the reduced space $\mathbb{R}^{\alpha}$ with $\alpha \leq m$, where:

- $n =$ number of observations;
- $m =$ number of covariates (input variables);
- $p =$ number of dependent variables (output variables);
- $\alpha =$ dimension of the reduced space in which X is projected.

Popular application of PLS

This technique is often used in **fault detection** and **isolation**. With PLS is possible to treat both regression and classification problems. The matrix X always contains the process variables (e.g. diameter and thickness of a gasket), while the matrix Y only (quantitative) quality variables (e.g. its mechanical seal) in the regression case, whereas in pattern classification the predicted variables are dummy variables (1 or 0) such as:

$$Y = \begin{bmatrix} 1 & \dots & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & 1 & \dots & 1 \end{bmatrix}^{\top} \tag{1}$$

where each column of Y corresponds to a fault class. The first n_j elements of column j are filled with a 1, which indicates that the first n_j rows of X are data from fault j. In this case PLS is called **discriminant**.

TO DO

TO DO

TO DO