Toulouse School of Economics

M2 EEE Machine Learning

Final Project

Andrew Boomer and Jacob Pichelmann

March 24, 2021

Table of Contents

Four dimension reduction devices: (1) Principal Component Analysis, (2) Ridge Regression, (3) Landweber Fridman (LF) regularization, (4) Partial least squares. Each involves a regularization or tuning parameter that is selected through genearlized cross validation (GCV).

Take two different data generating processes, (1) The eigenvalues of $\frac{X'X}{T}$ are bounded and decline to zero gradually. (2) Popular factor model with a finite number, r, of factors. Here, the r largest eigenvalues grow with N, while the remaining are bounded. In both cases, $\frac{X'X}{T}$ is ill-conditioned, which means the ratio of the largest to smallest eigenvalue diverges, and a regularization terms is needed to invert the matrix.

Data Generating Process:

Large Sample Case: $N = 200$ and $T = 500$ Small Sample Case: $N = 100$ and $T = 50$

$$\underbrace{x_t}_{(N\times 1)} = \underbrace{\Lambda}_{(N\times r)} \underbrace{F_t}_{(r\times 1)} + \underbrace{\xi_t}_{(N\times 1)}$$

$$\underbrace{y_t}_{(1\times 1)} = \underbrace{\theta'}_{(1\times r)} \underbrace{F_t}_{(r\times 1)} + \underbrace{\nu_t}_{(1\times 1)}$$

$$\underbrace{y}_{(T\times 1)} = \underbrace{F}_{(T\times r)} \underbrace{\theta}_{(r\times 1)} + \underbrace{\nu}_{(T\times 1)}$$

$$\underbrace{X}_{(T\times N)} = \underbrace{F}_{(T\times r)} \underbrace{\Lambda'}_{(r\times N)} + \underbrace{\xi}_{(T\times N)}$$

DGP 1 (Few Factors Structure): θ is the $(r \times 1)$ vector of ones, $r = 4$ and $r_{max} = r+10$

DGP 2 (Many Factors Structure): θ is the $(r \times 1)$ vector of ones, $r = 50$ and $r_{max} = min(N, \frac{T}{2})$

DGP 3 (Five Factors but only One Relevant): $\theta = (1, 0_{1\times 4})$, $r = 5$ and $r_{max} = min(r + 10, min(N, \frac{T}{2}))$

$F = [F_1, F_2']'$ and $F \times F' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$

$y = \hat{F}\theta + \nu$ where $\hat{F}$ is generated from X equation in DGP 3, and $\sigma_\nu = 0.1$

DGP 4 (x_t Has a Factor Structure but Unrelated to y_t):

θ is a vector of zeros with dimension $(r \times 1)$. $r = 5$, $r_{max} = r + 10$. $F \times F'$ is defined as in DGP 3.

DGP 5 (Eigenvalues Declining Slowly):

θ is an $(N \times 1)$ vector of ones. $r = N$, $r_{max} = min(N, \frac{T}{2}$.

$\Lambda = M \odot \xi$, with $\xi \sim (N \times N)$ matrix of $iidN(0, 1)$

$M \sim (N \times N) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \end{bmatrix}$

DGP 6 (Near Factor Model):

$\theta = 1$, $r = 1$, $r_{max} = r + 10$, $\Lambda' = \frac{1}{\sqrt{N}} 1_{r\times N}$

Estimation:

Set 1: Bai-Ng, PCA, PLS, Ridge, LF, LASSO Set 2: GCV, Mallows, AIC, BIC Set 3: Small Sample, Large Sample

Parameter Iteration (For Later) Simulation (Andy) Model Estimation (Jacob) Evaluation (Jacob) Output (Andy)

1	1	2	3	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1