HPC for numerical methods and data analysis

Fall Semester 2024

Prof. Laura Grigori

Assistant: Mariana Martínez Aguilar

Session 9 – November 12, 2024

Randomized low rank approximation

For $A \in \mathbb{R}^{n\times n}$ SPSD (symmetric positive semidefinite) we want to implement a randomized algorithm that approximates this matrix.

Exercise 1: Create test matrices

Build the following test matrices. Here $n = 10^3$ and $R \in \{5, 10, 20\}$.

Low-Rank and PSD Noise. Let A be in the following form:

$$A = \text{diag}(1, ..., 1, 0, ..., 0) + \xi n^{-1} W,$$

where there are R initial 1's followed by zeros, $W \in \mathbb{R}^{n\times n}$ has a Wishart distribution, $W \sim \text{WISHART}(n, n)$. That is $W = GG^\top$, where $G \in \mathbb{R}^{n\times n}$ is a standard normal matrix. The parameter ξ controls the signal-to-noise ratio. Consider three examples, $\xi = 10^{-4}$, $\xi = 10^{-2}$, and $\xi = 10^{-1}$.

Polynomial Decay. Let A be in the following form:

$$A = \text{diag}(1, ..., 1, 2^{-p}, 3^{-p}, ..., (n - R + 1)^{-p},$$

where there are R initial 1's. Let $p \in \{0.5, 1, 2\}$.

Exponential Decay. Let A be in the following form:

$$A = \text{diag}(1, ..., 1, 10^{-q}, 10^{-2q}, ..., 10^{-(n-R)q}),$$

where there are R initial 1's and the parameter $q > 0$ controls the rate of exponential decay. Let $q \in \{0.1, 0.25, 1\}$.

Exercise 2: Randomized Nyström

For $A \in \mathbb{R}^{n\times n}$ SPSD and a sketching $\Omega_1 \in \mathbb{R}^{n\times l}$, randomized Nyström approximation computes:

$$\tilde{A}_{\text{Nyst}} = (A\Omega_1)(\Omega_1^\top A\Omega_1)^\dagger(\Omega_1^\top A),$$

Algorithm 1 Randomized Nyström

Input: $A \in \mathbb{R}^{n\times n}$, $l \in \mathbb{N}$, sketching $\Omega_1 \in \mathbb{R}^{n\times l}$
Output: Approximation $\tilde{A}_{\text{Nyst}} = \hat{U}\Sigma^2\hat{U}^\top$
Compute $C = A\Omega_1$
Compute $B = \Omega_1^\top C$ and its Cholesky factorization $B = LL^\top$
Compute $Z = CL^{-\top}$
Compute the QR factorization of $Z = QR$
Compute the SVD factorization of $R = \tilde{U}\Sigma\tilde{V}^\top$
Compute $\hat{U} = Q\tilde{U}$
Output factorization $\tilde{A}_{\text{Nyst}} = \hat{U}\Sigma^2\hat{U}^\top$

where $(\cdot)^\dagger$ denotes the pseudoinverse. Consider the following algorithm:
Do the following:

a) Plot the singular values of the matrices built in exercise 1

b) Explain the idea behind Nyström factorization and possible problems with algorithm 1

c) Implement algorithm 1

d) For each of the test matrices, plot the singular values of B, compute the condition number of this matrix and explain why this might be a problem

e) Relate the condition number of B with computational difficulties when computing $Z = CL^{-\top}$

f) Propose a stable algorithm for computing Z in the test matrices

g) Plot the relative error, $\text{rel}(A, \tilde{A}_{\text{Nyst}})$

h) Comment on the relationship between the relative error with the condition number of A, the condition number of B and the computation of Z