ECS230 Applied Numerical Linear Algebra Fall 2024 2024-11-18

Homework 4 due 2024-11-26 at 23:59

Create a pdf file containing your report. Upload a zip file hw4.zip containing your program (source only) and your pdf report.

Polynomial fit using the Normal Equations

Write a C/C++ program that computes the coefficients of a polynomial that approximates a given data set in the least-square sense. The program should work for a polynomial of arbitrary degree and for an arbitrary number of data points. It should use the BLAS and LAPACK library functions for linear algebra operations. (Note: you will have to use the options **`-lblas -llapack`** when compiling your program)

1) Download the file **`data.dat`** from Canvas

2) Write a C or C++ program that will:
a) read the degree of the polynomial to be fitted from the first argument on the command line (**`argv[1]`**).
b) open the file **`data.dat`** and read the number of data points *n* from the first line of the file. Then read the *n* data points (x_i, y_i) *i=1,..,n* from the file.
c) compute and print the matrices *X*, X^TX and X^Ty that appear in the normal equations

$$X^T X b = X^T y$$

where

$$(X)_{ij} = (x_i)^{j-1}$$

The matrix product $A = X^T X$ should be computed using the **`dgemm`** function from the BLAS library. The right hand side $X^T y$ should be computed using the **`dgemv`** function.
d) compute the Cholesky factorization $A = LL^T$ using the function **`dpotrf`** of the LAPACK library, and print the matrix L.
e) solve the two triangular systems of equations required to compute the solution vector b containing the coefficients of the polynomial. The triangular systems of equations should be solved using the function **`dtrsm`** of the BLAS library.
f) write the original data and the polynomial fit on a plot file in a format suitable for visualization using e.g. the gnuplot program (or another plotting program of your choice).

3) Run the program using the data in the file "data.dat" and polynomials of degrees 3 to 9. Plot the data and the fit to the data. Observe the quality of the fit using the data written on the plot file. Comment on the quality of the fit obtained for different choices of degree. Include in your report the output of the program for the data in **`data.dat`** and the coefficients of the polynomial fit of degree 5 and of degree 9, as well as plots of the corresponding fits.
Discuss the quality of the fit for various choices of the degree. Why does the fit not work for a degree of 10?

4) Modify your program to include regularization by ridge regression. Discuss the effect of ridge parameter values 1, 10, 100 on the fits for degrees 1 and 5.