Comparison of gradient schemes

1 Gradient descent

Consider the following least square problem

$$\min_{x \in \mathbb{R}^n} \big\{ F(x) \stackrel{\text{def}}{=} \tfrac{1}{2}\|Ax - b\|^2 \big\}, \tag{1.1}$$

where $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$ is of the form

$$A = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & & \cdots & & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}_n .$$

Numerical setting Size of the problem $n = 201$, under this setting the condition number of the problem is roughly $\text{cnd} = 2.7 \times 10^8$. Vector b is a Gaussian random vector.

Algorithms Solve (1.1) with gradient descent, heavy-ball method and Nesterov's optimal scheme, compare the performance of these schemes. In terms of objective function value and $\|x_k - x_{k-1}\|$.

2 Proximal gradient descent

Consider the ℓ_1-regularised least square problem, which is also known as LASSO problem.

$$\min_{x \in \mathbb{R}^n} \mu\|x\|_1 + \tfrac{1}{2}\|Ax - b\|^2. \tag{2.1}$$

This is a typical problem that can be handled by proximal gradient descent method.

Data The `australian` data set from LIBSVM[1] is considered. See also the download link on course website.

Algorithms Solve (2.1) with proximal gradient descent, proximal version of heavy-ball method and FISTA scheme, compare the performance of these schemes. In terms of objective function value and $\|x_k - x_{k-1}\|$.

[1] https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/