A FAST ALGORITHM FOR LEAST SQUARES PROBLEMS

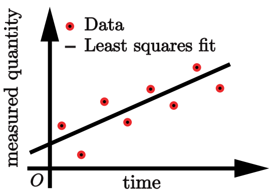
Inner-Outer Iteration Method for Least Squares Problems

Keiichi Morikuni (The Graduate University for Advanced Studies)
Ken Hayami (National Institute of Informatics, The Graduate University for Advanced Studies)
Xiaoke Cui (The University of Tokyo)

WHAT FOR ?

- Survey
- Control
- Signal processing
- Statistical processing
- Image processing
- Optimization
- etc.

e.g.



HOW ?

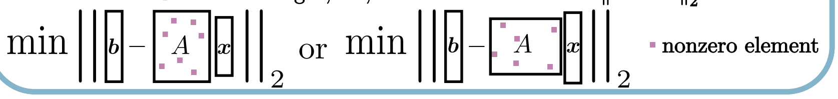
- Development of solvers for solving least squares problems
- Theoretical analysis of the property of the solvers
- Verification through numerical experiments

WHAT DO WE UNDERSTAND ?

Property of the solutions of least squares problems

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$$

where $A \in \mathbb{R}^{m\times n}$ is "large", i.e., find $\boldsymbol{x}$ which minimizes $\|\boldsymbol{b} - A\boldsymbol{x}\|_2$



HOW IS IT RELATED TO US ?

For example...

- To determine the 3-D structure of protein molecules
- To determine the inner constitution of the earth from the analysis of seismic waves

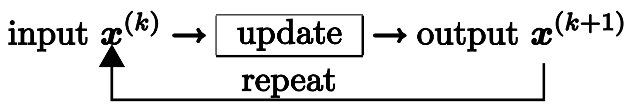
WHY DO WE STUDY ?

- To develop faster, more accurate, and more efficient methods for solving least squares problems
- To solve larger and larger, more and more ill-conditioned problems
- Theoretical understanding of the properties of the solvers

ITERATIVE METHOD

Iterative methods have advantages for large, sparse problems
The process is described as

1. Choose the initial solution $\boldsymbol{x}_0$
2. For $j = 1, 2, \ldots,$ Do
3. Update $\boldsymbol{x}_j$ with an algorithm
4. EndDo



We should choose or have to develop better algorithms to achieve faster, more robust, and more accurate convergence

Krylov subspace iterative methods for linear systems use

$$\mathcal{K}_m(A, \boldsymbol{r}_0) = \operatorname{span}\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{m-1}\boldsymbol{r}_0\} \text{ for } A \in \mathbb{R}^{n\times n} \text{ and initial residual } \boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$$

- The generalized minimal residual (GMRES) method minimizes $\|\boldsymbol{b} - A\boldsymbol{x}\|_2$ over the subspace
  $$\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} = \mathcal{K}_m(A, \boldsymbol{r}_0), \text{ where } (\boldsymbol{v}_{j+1}, \boldsymbol{v}_i) = 0 \text{ for } i = 1, 2, \ldots, j \text{ and } (\boldsymbol{v}_{j+1}, \boldsymbol{v}_{j+1}) = 1$$
- The conjugate gradient method for least squares problems (CGLS) applies the CG method to $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$

NUMERICAL EXPERIMENTS

Table 1: Information of the matrices

Problem	m	n	nnz	dens.[%]	rank	$\kappa(A)$
Maragal_6	21,251	10,144	537,694	0.25	8,331	2.91×10^6
Maragal_8	60,845	33,093	1,308,415	0.06	-	-

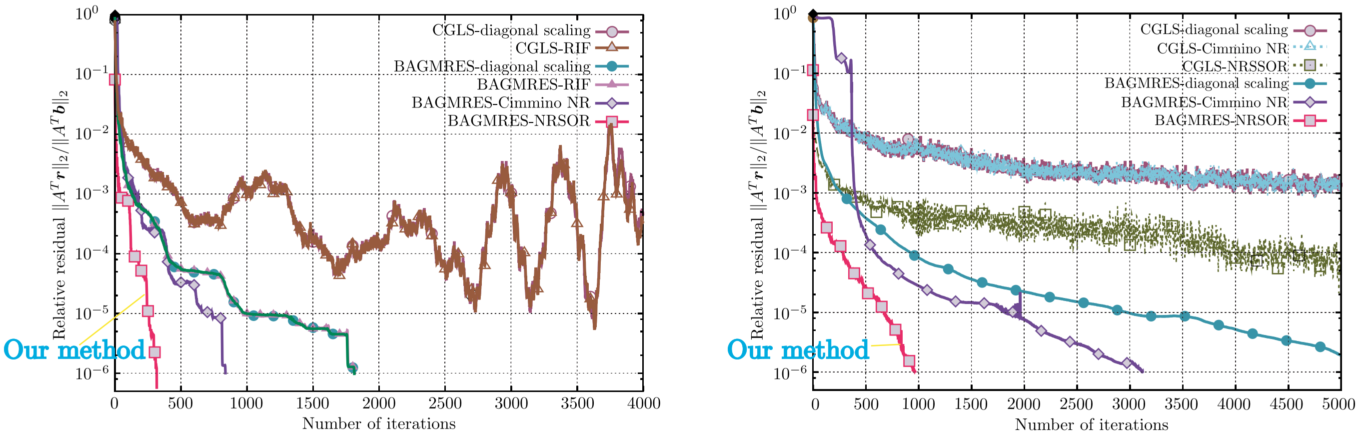


Figure 1: Convergence history of the relative residuals for Maragal_6 and _8