COMPARATIVE EFFICIENCIES OF THE LEAST SQUARES METHOD AND THE GRADIENT METHOD FOR OVERDETERMINED LINEAR SYSTEMS

Béla Finta
”Petru Maior” University of Tg. Mureş, Romania
e-mail: fintab@science.upm.ro

ABSTRACT

The purpose of this paper is to determine the arithmetic operations count for the least squares method and for the gradient method in the case of overdetermined linear systems. If the overdetermined linear system has great dimensions, then the arithmetic operations count with gradient method is less then with the least squares method.

Keywords: least squares method, gradient method, overdetermined linear system

1 Introduction

Let us consider the real matrix $A = (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}}$, and the real, transposed arrays $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ and $b = (b_1, b_2, \ldots, b_m)^T \in \mathbb{R}^m$, respectively. The linear system $A \cdot x = b$ is called overdetermined linear system, if $m > n$. Generally, the overdetermined linear system is incompatible, i.e. doesn't exist an array $x^* = (x_1^*, x_2^*, \ldots, x_n^*)^T \in \mathbb{R}^n$ such that $A \cdot x^* = b$. For this reason, instead of the classical solution x^*, we consider such array $\overline{x} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n) \in \mathbb{R}^n$ for which the function $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|A \cdot x - b\|_m^2$ takes the minimal value, where $\|\cdot\|_m$ means the Euclidean norm on the space $\mathbb{R}^m$. The array $\overline{x} \in \mathbb{R}^n$, which minimizes the function f, it is accepted like the solution of the overdetermined linear system $A \cdot x = b$ in the sense of the least squares method. We can observe that $f(x) \geq 0$ for all $x \in \mathbb{R}^n$ and the minimal point $\overline{x} \in \mathbb{R}^n$ verifies the following system with partial derivatives $\frac{\partial f}{\partial x_k}(\overline{x}) = 0$ for every $k = \overline{1,n}$. We calculate the partial derivatives and doing the corresponding calculus, we obtain that $\overline{x} \in \mathbb{R}^n$ is the solution of the linear system $(A^T \cdot A) \cdot x = A^T \cdot b$, which is a Cramer's type linear system with n equations and n unknowns, so $\overline{x} \in \mathbb{R}^n$ will be the classical solution of this Cramer's linear system. We also mention the following statement: if $\overline{x} \in \mathbb{R}^n$ is the classical solution of the linear system $(A^T \cdot A) \cdot x = A^T \cdot b$, i.e. $(A^T \cdot A) \cdot \overline{x} = A^T \cdot b$, then $f(\overline{x}) = \|A \cdot \overline{x} - b\|_m^2 \leq \|A \cdot x - b\|_m^2 = f(x)$ for all $x \in \mathbb{R}^n$. So $\overline{x} \in \mathbb{R}^n$ is the solution of the overdetermined linear system $A \cdot x = b$ in the sense of the least squares approache, see [1] and [3].

The linear system $(A^T \cdot A) \cdot x = A^T \cdot b$ can be solved by Cramer's rule from theory of determinants for $n \in \mathbb{N}^*$ small natural numbers, but in other cases, for $n \in \mathbb{N}^*$ great natural numbers we can use numerical methods of linear algebra. We mention here the Gauss elimination method, when the Cramer type linear system is reduced to an upper triangular linear system, [1] or [3]. Analogously, we can reduce the Cramer type linear system to a lower triangular linear system, too. When we reduce the Cramer type linear system to a diagonal linear system then we use the Gauss-Jordan elimination procedure, [3]. In [2] we presented the gradient method for overdetermined linear systems $A \cdot x = b$.

Let us choose $x^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in \mathbb{R}^n$ the start point for the gradient method, and we suppose that we determined the point $x^k = (x_1^k, x_2^k, \ldots, x_n^k) \in \mathbb{R}^n$ and we want to find the next point $x^{k+1} = (x_1^{k+1}, x_2^{k+1}, \ldots, x_n^{k+1}) \in \mathbb{R}^n$ with the gradient method. In [2] we obtained the formula: $x^{k+1} = x^k - \alpha_k \cdot A^T \cdot (A \cdot x^k - b)$ with

$$\alpha_k = \frac{< A \cdot x^k - b, A \cdot A^T \cdot (A \cdot x^k - b) >_m}{< A \cdot A^T \cdot (A \cdot x^k - b), A \cdot A^T \cdot (A \cdot x^k - b) >_m},$$

where $< \cdot, \cdot >_m$ is the Euclidean scalar product on $\mathbb{R}^m$.

2 Main part

The purpose of this paper is to compare the arithmetic operations count for the least squares method and for the gradient method in the case of overdetermined linear systems. In order to obtain $A^T \cdot A$ we must multiply any row of A^T with any column of A, so we need m multiplication and $m-1$ addition and this is repeted n^2 times, which means $n^2 \cdot m$ multiplication and $n^2 \cdot (m-1) = n^2 \cdot m - n^2$ addition. In order to obtain $A^T \cdot b$ we must multiply a row of A^T with the column matrix b, so we need m multiplication and $m-1$ addition and for any row of A^T it is repeted n times, which means $n \cdot m$ multiplication and $n \cdot (m-1) = n \cdot m - n$ addition. Hence we can form the linear system $(A^T \cdot A) \cdot x = A^T \cdot b$ with $n^2 \cdot m + n \cdot m$ multiplication and $n^2 \cdot m - n^2 + n \cdot m - n$ addition. According to [3], if we solve the Cramer type linear system $(A^T \cdot A) \cdot x = A^T \cdot b$ with Gauss elimination method we use $\frac{4}{3} \cdot n^3 - \frac{1}{3} \cdot n$ multiplications and divisions and $\frac{4}{3} \cdot n^3 - \frac{3}{2} \cdot n^2 + \frac{1}{6} \cdot n$ additions and subtractions, and if we solve it with Gauss-Jordan elimination method, then we use $\frac{3}{2} \cdot n^3 - \frac{1}{2} \cdot n$ multiplications and divisions and $\frac{3}{2} \cdot n^3 - 2 \cdot n^2 + \frac{1}{2} \cdot n$ additions and subtractions, respectively. Hence to solve the overdetermined linear system by the least squares method involving the Gauss elimination method we execute $n^2 \cdot m + n \cdot m + \frac{4}{3} \cdot n^3 - \frac{1}{3} \cdot n$ multiplications and divisions and $n^2 \cdot m - n^2 + n \cdot m - n + \frac{4}{3} \cdot n^3 - \frac{3}{2} \cdot n^2 + \frac{1}{6} \cdot n = n^2 \cdot m + n \cdot m + \frac{4}{3} \cdot n^3 - \frac{5}{2} \cdot n^2 - \frac{5}{6} \cdot n$ additions and subtractions, and involving the Gauss-Jordan elimination method we make $n^2 \cdot m + n \cdot m + \frac{3}{2} \cdot n^3 - \frac{1}{2} \cdot n$ multiplications and divisions and $n^2 \cdot m - n^2 + n \cdot m - n + \frac{3}{2} \cdot n^3 - 2 \cdot n^2 + \frac{1}{2} \cdot n = n^2 \cdot m + n \cdot m + \frac{3}{2} \cdot n^3 - 3 \cdot n^2 - \frac{1}{2} \cdot n$ additions and subtractions, respectively.

Next we calculate the number of arithmetical operations necessary in the case of gradient method in order to solve the overdetermined linear system $A \cdot x = b$. We remember the formulas at the step $k+1 : x^{k+1} = x^k - \alpha_k \cdot A^T \cdot (A \cdot x^k - b)$, where

$$\alpha_k = \frac{< A \cdot x^k - b, A \cdot A^T \cdot (A \cdot x^k - b) >_m}{< A \cdot A^T \cdot (A \cdot x^k - b), A \cdot A^T \cdot (A \cdot x^k - b) >_m}.$$

If a row of the matrix A we multiply with the column array $x^k \in \mathbb{R}^n$ we execute n multiplications and $n-1$ additions. If we consider all the m rows of A then for $A \cdot x^k$ we make $m \cdot n$ multiplications and $m \cdot (n-1)$ additions. To realize $A \cdot x^k - b$ we do m subtractions. If a row of the matrix A^T we multiply with the column vector $A \cdot x^k - b$ we have m multiplications and $m-1$ additions. If we take all the n rows of A^T then for $A^T \cdot (A \cdot x^k - b)$ we execute $n \cdot m$ multiplications and $n \cdot (m-1)$ additions. If a row of the matrix A we multiply with the column vector $A^T \cdot (A \cdot x^k - b)$ we do n multiplications and $n-1$ additions. If we take all the m rows of A then for $A \cdot A^T \cdot (A \cdot x^k - b)$ we do $m \cdot n$ multiplications and $m \cdot (n-1)$ additions. In order to obtain the scalar products $< A \cdot x^k - b,\ A \cdot A^T \cdot (A \cdot x^k - b) >_m$ and $< A \cdot A^T \cdot (A \cdot x^k - b), A \cdot A^T \cdot (A \cdot x^k - b) >_m$ we execute m multiplications and $m-1$ additions, respectively. In order to obtain the value α_k, we make one division, and to get $\alpha_k \cdot A^T \cdot (A \cdot x^k - b)$ we do n multiplication, because the scalar α_k multiply the column vector $A^T \cdot (A \cdot x^k - b)$ and after we calculate x^{k+1}, using the vector subtraction $x^k - \alpha_k \cdot A^T \cdot (A \cdot x^k - b)$, which means n subtractions. Totally we have $m \cdot n + n \cdot m + m \cdot n + m + m + 1 + n = 3 \cdot m \cdot n + 2 \cdot m + n + 1$ multiplications and divisions and $m \cdot (n-1) + m + n \cdot (m-1) + m \cdot (n-1) + (m-1) + (m-1) + n = 3 \cdot m \cdot n + m - 2$ additions and subtractions.

Now, if with the gradient method we make k steps, then we do totally $k \cdot [3 \cdot m \cdot n + 2 \cdot m + n + 1]$ multiplications and divisions and $k \cdot [3 \cdot m \cdot n + m - 2]$ additions and subtractions.

3 Conclusions

Let $m, n \in \mathbb{N}^*$ be great natural numbers with $m > n$, so we suppose that the overdetermined linear system has great dimensions. If $k \in \mathbb{N}^*$, the number of iteration steps verifies $k < \frac{n}{3}$, then $k \cdot [3 \cdot m \cdot n + 2 \cdot m + n + 1] < n^2 \cdot m + n \cdot m + \frac{4}{3} \cdot n^3 - \frac{1}{3} \cdot n$, $k \cdot [3 \cdot m \cdot n + 2 \cdot m + n + 1] < n^2 \cdot m + n \cdot m + \frac{3}{2} \cdot n^3 - \frac{1}{2} \cdot n$, and $k \cdot [3 \cdot m \cdot n + m - 2] < n^2 \cdot m + n \cdot m + \frac{4}{3} \cdot n^3 - \frac{5}{2} \cdot n^2 - \frac{5}{6} \cdot n$, $k \cdot [3 \cdot m \cdot n + m - 2] < n^2 \cdot m + n \cdot m + \frac{3}{2} \cdot n^3 - 3 \cdot n^2 - \frac{1}{2} \cdot n$. So we can conclude, that for great overdetermined linear systems is better to use the gradient method instead of the least squares method, because we can reduce the arithmetic operations count.

References

[1] B. Finta, Analiză numerică, Editura Universităţii „Petru Maior", Tg. Mureş, 2004.

[2] B. Finta, The Gradient Method for Overdetermined Linear Systems, Scientific Bulletin of the „Petru Maior" University of Tg. Mureş, Vol.7 (XXIV), no. 1, 2010, ISSN 1841-9267.

[3] S. S. Rao, Applied Numerical Methods for Engineers and Scientists, Prentice Hall, Upper Saddle River, New Jersey, 2002.