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

FINTA Béla, Associate Professor Ph.D.,
"Petru Maior" University of Tîrgu Mureş

Abstract: The purpose of this paper is to make the comparative efficiencies of the least squares method and the gradient method for weighted overdetermined linear systems.

Keywords: least squares method, gradient method, overdetermined linear system, weighted 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 $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ and $b = (b_1, b_2, \dots, b_m)^T \in \mathbb{R}^m$, 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 $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T \in \mathbb{R}^n$ such that $A \cdot x^* = b$. It is well known to obtain the solution of the overdetermined linear system using the least squares method, see for example [1] and [2]. In [3] we solved the overdetermined linear system $A \cdot x = b$ using the gradient method and in [4] we did the comparative efficiencies of the least squares method and the gradient method for the overdetermined linear system $A \cdot x =$ b.

Let us consider the weight array $p = (p_1, p_2, \dots, p_m)^T \in \mathbb{R}^m$, where $p_i > 0$ for every $i = \overline{1,m}$. We denote by $\sqrt{p} = (\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_m}) \in \mathbb{R}^m$ and for the vectors $x = (x_1, x_2, \dots, x_m)^T \in \mathbb{R}^m$ and $y = (y_1, y_2, \dots, y_m)^T \in R^m$ we introduce the following product $x \otimes y = (x_1 \cdot y_1, x_2 \cdot y_2, \dots, x_m \cdot y_m) \in R^m$. Let us take the weighted matrix $A_p = (\sqrt{p_i} \cdot a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}$, and the weighted vector $b_p = \sqrt{p} \otimes b \in \mathbb{R}^m$. So to the overdetermined linear system $A \cdot x = b$ we can attach the weighted overdetermined linear system $A_p \cdot x = b_p$. It is immediatly that these two linear systems are equivalent. This means that, generally, the weighted overdetermined linear system is incompatible, i.e. doesn't exist an array $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T \in R^n$ such that $A_p \cdot x^* = b_p$. For this reason, instead of the classical solution x^*, we consider such array $\bar{x}' = (\bar{x}_1', \bar{x}_2', \dots, \bar{x}_n') \in \mathbb{R}^n$ for which the function $f' : \mathbb{R}^n \to \mathbb{R}, f'(x) = \|A_p \cdot x - b_p\|_m^2$ takes the minimal value, where $\|\cdot\|_m$ means the Euclidean norm on the space $\mathbb{R}^m$. In [5] we showed the least squares method for weighted overdetermined linear systems and in [6] we presented the gradient method for weighted overdetermined linear systems, too.

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 weighted overdetermined linear systems $A_p \cdot x = b_p$. In order to obtain $A_p^T \cdot A_p$ we must multiply any row of A_p^T with any column of A_p, so we need m multiplications and $m - 1$ additions and this is repeted n^2 times, which means $n^2 \cdot m$ multiplications and $n^2 \cdot (m-1) = n^2 \cdot m - n^2$ additions. In order to obtain $A_p^T \cdot b_p$ we must multiply a row of A_p^T with the column matrix b_p, so we need m multiplications and $m - 1$ additions and for any row of A_p^T it is repeted n times, which means $n \cdot m$ multiplications and $n \cdot (m-1) = n \cdot m - n$ additions. Hence we can form the linear system $\left(A_p^T \cdot A_p\right) \cdot x = A_p^T \cdot b_p$ with $n^2 \cdot m + n \cdot m$ multiplications and $n^2 \cdot m - n^2 + n \cdot m - n$ additions. According to [2], if we solve the Cramer type linear system $\left(A_p^T \cdot A_p\right) \cdot x = A_p^T \cdot b_p$ 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 weighted 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 weighted overdetermined linear system $A_p \cdot x = b_p$. We remember the formulas at the step $k+1$:

$$x^{k+1} = x^k - \alpha'_k \cdot A_p^T \cdot (A_p \cdot x^k - b_p),$$

where $\alpha'_k = \frac{N_k}{D_k}$, with nominator

$$N_k = < A_p \cdot x^k - b_p, A_p \cdot A_p^T \cdot \left(A_p \cdot x^k - b_p\right) >_m$$

and denominator

$$D_k = < A_p \cdot A_p^T \cdot \left(A_p \cdot x^k - b_p\right), A_p \cdot A_p^T \cdot \left(A_p \cdot x^k - b_p\right) >_m.$$

If a row of the matrix A_p we multiply with the column array $x^k \in R^n$ we execute n multiplications and $n - 1$ additions. If we consider all the m rows of A_p then for $A_p \cdot x^k$ we make $m \cdot n$ multiplications and $m \cdot (n-1)$ additions. To realize $A_p \cdot x^k - b_p$ we do m subtractions. If a row of the matrix A_p^T we multiply with the column vector $A_p \cdot x^k - b_p$ we have m multiplications and $m-1$ additions. If we take all the n rows of A_p^T then for $A_p^T \cdot (A_p \cdot x^k - b_p)$ we execute $n \cdot m$ multiplications and $n \cdot (m-1)$ additions. If a row of the matrix A_p we multiply with the column vector $A_p^T \cdot (A_p \cdot x^k - b_p)$ we do n multiplications and $n-1$ additions. If we take all the m rows of A_p then for $A_p \cdot A_p^T \cdot (A_p \cdot x^k - b_p)$ we do $m \cdot n$ multiplications and $m \cdot (n-1)$ additions. In order to obtain the scalar products $< A_p \cdot x^k - b_p, A_p \cdot A_p^T \cdot (A_p \cdot x^k - b_p) >_m$ and $< A_p \cdot A_p^T \cdot (A_p \cdot x^k - b_p), A_p \cdot A_p^T \cdot (A_p \cdot x^k - b_p) >_m$ we execute m multiplications and $m-1$ additions, respectively. In order to obtain the value α'_k, we make one division, and to get $\alpha'_k \cdot A_p^T \cdot (A_p \cdot x^k - b_p)$ we do n multiplications, because the scalar α'_k multiply the column vector $A_p^T \cdot (A_p \cdot x^k - b_p)$ and after we calculate x^{k+1}, using the vector subtraction $x^k - \alpha'_k \cdot A_p^T \cdot (A_p \cdot x^k - b_p)$, 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.

Consequence 1. For $p_1 = p_2 = \cdots = p_m = 1$ from the above we reobtain the result presented in [4].

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

$$k \cdot [3 \cdot m \cdot n + 2 \cdot m + n + 1] < n^2 \cdot m + n \cdot m + \frac{4}{3} \cdot n^3 - \frac{1}{3} \cdot n,$$

$$k \cdot [3 \cdot m \cdot n + 2 \cdot m + n + 1] < n^2 \cdot m + n \cdot m + \frac{3}{2} \cdot n^3 - \frac{1}{2} \cdot n, \text{ and}$$

$$k \cdot [3 \cdot m \cdot n + m - 2] < n^2 \cdot m + n \cdot m + \frac{4}{3} \cdot n^3 - \frac{5}{2} \cdot n^2 - \frac{5}{6} \cdot n,$$

$$k \cdot [3 \cdot m \cdot n + m - 2] < n^2 \cdot m + n \cdot m + \frac{3}{2} \cdot n^3 - 3 \cdot n^2 - \frac{1}{2} \cdot n.$$

So we can conclude, that for great weighted 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. S. S. Rao, Applied Numerical Methods for Engineers and Scientists, Prentice Hall, Upper Saddle River, New Jersey, 2002.
3. 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, pp. 24-26.
4. B. Finta, Comparative Efficiencies of the Least Squares Method and the Gradient Method for Overdetermined Linear Systems, Scientific Bulletin of the "Petru Maior" University of Tg. Mureş, Vol.7 (XXIV), no. 2, 2010, ISSN 1841-9267, pp. 35-36.
5. B. Finta, Weighted overdetermined linear systems, Scientific Bulletin of the "Petru Maior" University of Târgu Mureş, Vol. 10 (XXVII), no. 1, 2013, ISSN-L 1841-9267, pp.47-48.
6. B. Finta, The Gradient Method for Weighted Overdetermined Linear Systems, Inter-Eng Conference, "Petru Maior" University of Târgu Mureş, 2013, (submitted).