ИВМ СО РАН
Отдел вычислительной математики

Киреев И. В.

РЕШЕНИЕ ЛИНЕЙНОЙ ЗАДАЧИ НАИМЕНЬШИХ КВАДРАТОВ МЕТОДОМ СОПРЯЖЁННЫХ НАПРАВЛЕНИЙ

Красноярск 2011

1. Линейная задача метода наименьших квадратов.

Least Squares Problem

По заданным действительным $m \times n$**-матрице** A **ранга** $k \le \min(m,n)$ **и** m**-векторе** b **найти действительный** n**-вектор** x**, минимизирующий евклидову длину вектора** $A\boldsymbol{x} - \boldsymbol{b}$ [1],

$$\boldsymbol{x} \in \underset{\boldsymbol{y}\in\mathbb{R}^n}{\operatorname{argmin}} \|A\boldsymbol{y} - \boldsymbol{b}\|_m^2, \qquad (1)$$

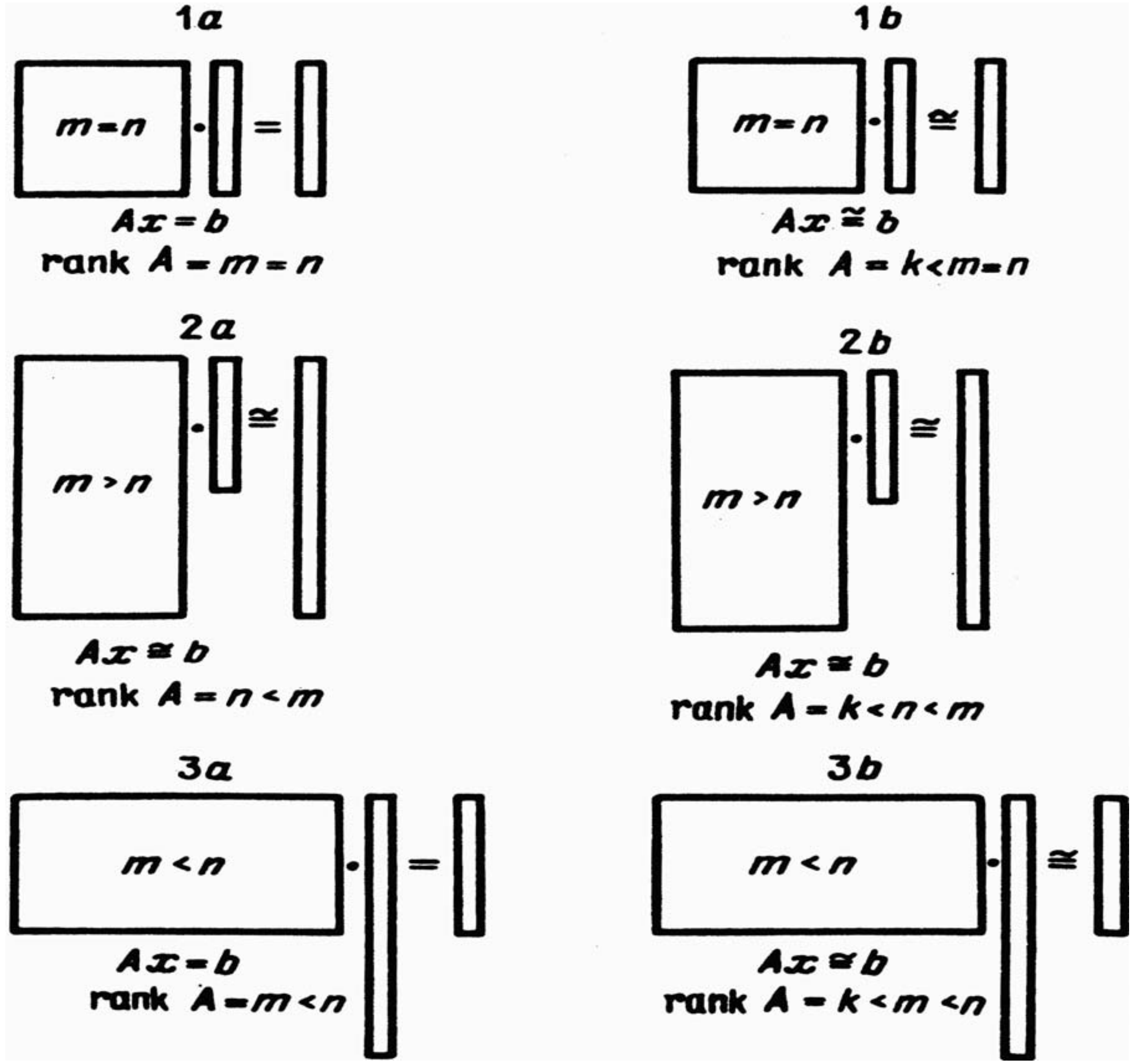
здесь $\quad \|\boldsymbol{y}\|_n^2 = y_1^2 + y_2^2 + \ldots + y_n^2,$

$$y_1, y_2, \ldots, y_n \text{ – } \textbf{координаты } \boldsymbol{y} \in \mathbb{R}^n;$$

обозначают

$$A\boldsymbol{x} \cong \boldsymbol{b}; \boldsymbol{b} \in \mathbb{R}^m, \boldsymbol{x} \in \mathbb{R}^n, k = \textbf{rang} A \le \textbf{min}(m,n). \qquad (2)$$

2. 6 случаев [1] линейной LSP



3. Система линейных уравнений LSP(2)

$$A^T(A\boldsymbol{x} - \boldsymbol{b}) = \mathbf{0}. \tag{3}$$

где

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \boldsymbol{b} = \begin{pmatrix} b_1 \\ \cdots \\ b_m \end{pmatrix} \in \mathbb{R}^m.$$

4. Методы сопряжённых градиентов

Conjugate Gradient Methods

Опр.[2] *Крыловское подпространство* $\mathcal{K}_k(B, \boldsymbol{c})$, *порождённое* $n \times n$*-матрицей* B *и вектором* $\boldsymbol{c} \in \mathbb{R}^n$ – *это линейная оболочка* $\mathcal{K}_k(B, \boldsymbol{c}) = \mathrm{span}\left\{\boldsymbol{c}, B\boldsymbol{c}, \ldots, B^{k-1}\boldsymbol{c}\right\}$.

Приближение $\boldsymbol{x}^{(k)}$, полученное на k-ой итерации CGLS, минимизирует функционал ошибки

$$E_\mu(\boldsymbol{y}) = (\boldsymbol{x} - \boldsymbol{y})^T \left(A^T A\right)^\mu (\boldsymbol{x} - \boldsymbol{y}), \qquad (4)$$

на множестве $\boldsymbol{y} \in \boldsymbol{x}^{(0)} + \mathcal{K}_k\left(A^T A, \boldsymbol{s}^{(0)}\right)$, где

$$\boldsymbol{s}^{(0)} = A^T \left(\boldsymbol{b} - A\boldsymbol{x}^{(0)}\right),$$

а $\mu = 0, 1, 2$.

5. Функционалы E_μ, $\mu = 0, 1, 2$

$$\mu = 0 \Rightarrow \ E_0(\boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|_n^2;$$

$$\mu = 1 \Rightarrow \ \mathbf{E_1(\boldsymbol{y}) = \|(\boldsymbol{b} - A\boldsymbol{x}) - (\boldsymbol{b} - A\boldsymbol{y})\|_m^2};$$

$$\mu = 2 \Rightarrow \ E_2(\boldsymbol{y}) = \|A^T(\boldsymbol{b} - A\boldsymbol{r}) - A^T(\boldsymbol{b} - A\boldsymbol{y})\|_n^2.$$

Теорема 1. *[3]*

$$E_\mu\left(\boldsymbol{x}^{(k)}\right) < 2\left(\frac{æ(A) - 1}{æ(A) + 1}\right)^k E_\mu\left(\boldsymbol{x}^{(0)}\right),$$

где $æ(A)$ **- число обусловленности матрицы** A.

6. Алгоритм CGLS ($\mu = 1$) [2]

Initialize

$$\boldsymbol{r}^{(0)} = \boldsymbol{b} - A\boldsymbol{x}^{(0)}, \boldsymbol{p}^{(0)} = \boldsymbol{s}^{(0)} = A^T\boldsymbol{r}^{(0)}, \gamma_0 = \|\boldsymbol{s}^{(0)}\|_m^2;$$

for $k = 0, 1, 2, \ldots$ **while** $\gamma_k > tol$ **compute**

$$\begin{aligned}
\boldsymbol{q}^{(k)} &= A\boldsymbol{p}^{(k)}, \\
\alpha_k &= \gamma_k / \|\boldsymbol{q}^{(k)}\|_m^2, \\
\boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}^{(k)}, \\
\boldsymbol{r}^{(k+1)} &= \boldsymbol{r}^{(k)} - \alpha_k \boldsymbol{p}^{(k)}, \\
\boldsymbol{s}^{(k+1)} &= A^T \boldsymbol{r}^{(k+1)}, \\
\gamma_{k+1} &= \|\boldsymbol{s}^{(k+1)}\|_n^2, \\
\beta_k &= \gamma_{k+1}/\gamma_k, \\
\boldsymbol{p}^{(k+1)} &= \boldsymbol{s}^{(k+1)} + \beta_k \boldsymbol{p}^{(k)}.
\end{aligned}$$

7. Точность алгоритма CGLS[4-5]

$$\frac{\|\boldsymbol{b} - A\boldsymbol{x}^{(k)} - \boldsymbol{r}^{(k)}\|_m}{\|A\|(\|\boldsymbol{x}\|_n + \|\boldsymbol{x}^{(0)}\|_n)} \leq$$

$$\leq \varepsilon_c \left((k + 1 + c) + k(10 + 2c) \max_{j \leq k} \frac{\|\boldsymbol{x}^{(j)}\|_n}{\|\boldsymbol{x}\|_n + \|\boldsymbol{x}^{(0)}\|_n} \right)$$

Здесь ε_c – машинная точность, c – связано с точностью вычисления произведения матрицы на вектор и обычно $c = m\sqrt{n}$.

8. Устойчивый алгоритм CGLS

Initialize

$$\boldsymbol{p}^{(0)} = \boldsymbol{s}^{(0)} = A^T\left(\boldsymbol{b} - A\boldsymbol{x}^{(0)}\right), \gamma_0 = \|\boldsymbol{s}^{(0)}\|_m^2;$$

for $k = 0, 1, 2, \ldots$ **while** $\gamma_k > tol$ **real compute**

$$\begin{aligned}
\boldsymbol{q}^{(k)} &= A\boldsymbol{p}^{(k)}, \\
\alpha_k &= \gamma_k / \|\boldsymbol{q}^{(k)}\|_m^2, \\
\boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}^{(k)}, \\
\boldsymbol{s}^{(k+1)} &= \boldsymbol{s}^{(k)} - \alpha_k\left(A^T\boldsymbol{q}^{(k)}\right), \\
\gamma_{k+1} &= \|\boldsymbol{s}^{(k+1)}\|_n^2, \\
\beta_k &= \gamma_{k+1}/\gamma_k, \\
\boldsymbol{p}^{(k+1)} &= \boldsymbol{s}^{(k+1)} + \beta_k \boldsymbol{p}^{(k)}.
\end{aligned}$$

9. Предобуславливание в LSP
Preconditioned problem

$$\boldsymbol{x} = S^{-1}\boldsymbol{y} \Rightarrow AS^{-1}\boldsymbol{y} \cong \boldsymbol{b}; \boldsymbol{b} \in \mathbb{R}^m, \boldsymbol{y} \in \mathbb{R}^n. \qquad (5)$$

Условия стационарности в PCLSP

$$\boldsymbol{x} = S^{-1}\boldsymbol{y} \Rightarrow S^{-T}A^{T}\left(AS\boldsymbol{x} - \boldsymbol{b}\right) = \boldsymbol{0}.$$

CGLS с предобуславливанием:

Теорема 2. *[2]*

$$\left\|\boldsymbol{r} - \boldsymbol{r}^{(k)}\right\|_m < 2\left(\frac{æ(AS^{-1}) - 1}{æ(AS^{-1}) + 1}\right)^k \left\|\boldsymbol{r} - \boldsymbol{r}^{(0)}\right\|_m,$$

где $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$, $\boldsymbol{r}^{(k)} = \boldsymbol{b} - A\boldsymbol{x}^{(k)}$, $æ(AS^{-1})$ **- число обусловленности матрицы** AS^{-1}.

10. Алгоритм с предобуславливанием PCCGLS

Initialize

$$\boldsymbol{r}^{(0)} = \boldsymbol{b} - A\boldsymbol{x}^{(0)}, \boldsymbol{p}^{(0)} = \boldsymbol{s}^{(0)} = S^{-T}\left(A^T\boldsymbol{r}^{(0)}\right), \gamma_0 = \|\boldsymbol{s}^{(0)}\|_m^2;$$

for $k = 0, 1, 2, \ldots$ **while** $\gamma_k > tol$ **compute**

$$\begin{aligned}
\boldsymbol{t}^{(k)} &= S^{-1}\boldsymbol{p}^{(k)}, \\
\boldsymbol{q}^{(k)} &= A\boldsymbol{t}^{(k)}, \\
\alpha_k &= \gamma_k / \|\boldsymbol{q}^{(k)}\|_m^2, \\
\boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{t}^{(k)}, \\
\boldsymbol{r}^{(k+1)} &= \boldsymbol{r}^{(k)} - \alpha_k \boldsymbol{q}^{(k)}, \\
\boldsymbol{s}^{(k+1)} &= S^{-T}\left(A^T\boldsymbol{r}^{(k+1)}\right), \\
\gamma_{k+1} &= \|\boldsymbol{s}^{(k+1)}\|_n^2, \\
\beta_k &= \gamma_{k+1}/\gamma_k, \\
\boldsymbol{p}^{(k+1)} &= \boldsymbol{s}^{(k+1)} + \beta_k \boldsymbol{p}^{(k)}.
\end{aligned}$$

11. Литература

1. **Лоусон Ч., Хенсон Р.** *Численное решение задач метода наименьших квадратов. Пер. с англ. – М.: Наука. Гл. ред. физ.-мат. лит., 1986. — 232 с.*

2. **Å. Björck.** *Numerical Methods for Least Squares Problems. – SIAM, Philadelphia,1996.-427 p.*

3. **Hestenes M. R. and Stiefel E.** *Methods of conjugate gradients for solving linear system, J. Res. Nat. Bur. Standards., B49 (1952), pp. 409-436.*

4. **Å. Björck, T. Elfving and Z. Strakos** *Stability of conjugate gradient-type methods for linear least squares problems, Technical Report LiTH-MAT-R-1995-26, Department of Mathematics, Linköping University, (1994)*

5. **A. Greenbaum** *Estimating the attainable accuracy of recursively computed residual methods, Technical Report TR-95-1515, Department of Computer Science, Cornell University, Ithaca, NY, (May 1995).*