

Iterative 5

王胤雅

25114020018

yinyawang25@m.fudan.edu.cn

2025 年 11 月 6 日

PROBLEM 1 Let $\phi_j(t)$ and $\pi_j(t)$ be the residual polynomial and the conjugate direction polynomial, respectively, for the BCG algorithm, as defined in Section 7.4.1. Let $\psi_j(t)$ be any other polynomial sequence which is defined from the recurrence

$$\begin{aligned}\psi_0(t) &= 1, & \psi_1(t) &= (1 - \xi_0 t)\psi_0(t), \\ \psi_{j+1}(t) &= (1 + \eta_j - \xi_j t)\psi_j(t) - \eta_j \psi_{j-1}(t).\end{aligned}$$

1. Show that the polynomials ψ_j are consistent, i.e.,

$$\psi_j(0) = 1 \quad \text{for all } j \geq 0.$$

2. Show the following relations:

$$\psi_{j+1}\phi_{j+1} = \psi_j\phi_{j+1} - \eta_j(\psi_{j-1} - \psi_j)\phi_{j+1} - \xi_j t \psi_j \phi_{j+1}, \quad (1)$$

$$\psi_j\phi_{j+1} = \psi_j\phi_j - \alpha_j t \psi_j \pi_j, \quad (2)$$

$$(\psi_{j-1} - \psi_j)\phi_{j+1} = \psi_{j-1}\phi_j - \psi_j\phi_{j+1} - \alpha_j t \psi_{j-1} \pi_j, \quad (3)$$

$$\psi_{j+1}\pi_{j+1} = \psi_{j+1}\phi_{j+1} - \beta_j \eta_j \psi_{j-1} \pi_j + \beta_j (1 + \eta_j) \psi_j \pi_j - \beta_j \xi_j t \psi_j \pi_j, \quad (4)$$

$$\psi_j \pi_{j+1} = \psi_j \phi_{j+1} + \beta_j \psi_j \pi_j. \quad (5)$$

3. Defining

$$t_j = \psi_j(A)\phi_{j+1}(A)r_0, \quad y_j = (\psi_{j-1}(A) - \psi_j(A))\phi_{j+1}(A)r_0,$$

$$p_j = \psi_j(A)\pi_j(A)r_0, \quad s_j = \psi_{j-1}(A)\pi_j(A)r_0,$$

show how the recurrence relations of the previous question translate for these vectors.

4. Find a formula that allows one to update the approximation x_{j+1} from the vectors x_{j-1} , x_j and t_j , p_j , y_j , s_j defined above.
5. Proceeding as in BICGSTAB, find formulas for generating the BCG coefficients α_j and β_j from the vectors defined in the previous question.

SOLUTION. 1. 将递推式在 $t = 0$ 处代入:

$$\psi_{j+1}(0) = (1 + \eta_j - 0)\psi_j(0) - \eta_j\psi_{j-1}(0) = (1 + \eta_j)\psi_j(0) - \eta_j\psi_{j-1}(0).$$

初始值 $\psi_0(0) = 1$, $\psi_1(0) = (1 - \xi_0 \cdot 0)\psi_0(0) = 1$ 。假设 $\psi_{j-1}(0) = \psi_j(0) = 1$, 则 $\psi_{j+1}(0) = (1 + \eta_j) - \eta_j = 1$ 。由归纳法, $\psi_j(0) = 1, \forall j \geq 0$ 。

2. 由于 ϕ_j, π_j 满足以下关系:

$$\phi_{j+1}(t) = \phi_j(t) - \alpha_j t \pi_j(t), \quad (1)$$

$$\pi_{j+1}(t) = \pi_j(t) + \beta_j \pi_j(t). \quad (2)$$

以及题中给的 ψ 递推

$$\psi_{j+1} = (1 + \eta_j - \xi_j t)\psi_j - \eta_j\psi_{j-1}.$$

可以得出以上关系式:

- $\psi_{j+1}\phi_{j+1} = (1 + \eta_j - \xi_j t)\psi_j\phi_{j+1} - \eta_j\psi_{j-1}\phi_{j+1} = \psi_{j+1}\phi_{j+1} = \psi_j\phi_{j+1} - \eta_j(\psi_{j-1} - \psi_j)\phi_{j+1} - \xi_j t \psi_j \phi_{j+1}$
- $\psi_j\phi_{j+1} = \psi(\phi_j - \alpha_j t \pi_j) = \psi_j\phi_j - \alpha_j t \psi_j \pi_j$
- $(\psi_{j-1} - \psi_j)\phi_{j+1} = \psi_{j-1}\phi_{j+1} - \psi_j\phi_{j+1} = -\psi_j\phi_{j+1} + \psi_{j-1}(\alpha_j t \pi_j + \phi_j) = \psi_{j-1}\phi_j - \phi_{j+1}\psi_j - \alpha_j t \psi_{j-1}\pi_j$
- $\psi_{j+1}\pi_{j+1} = \psi_{j+1}(\phi_{j+1} + \beta_j \pi_j) = \psi_{j+1}\phi_{j+1} + \psi_{j+1}\beta_j \pi_j = \psi_{j+1}\phi_{j+1} + \beta_j \pi_j((1 + \eta_j - \xi_j t)\psi_j(t) - \eta_j\psi_{j-1}(t)) = \psi_{j+1}\phi_{j+1} - \beta_j \eta_j \psi_{j-1}\pi_j + \beta_j(1 + \eta_j)\psi_j\pi_j - \beta_j \xi_j t \psi_j \pi_j$
- $\phi_j \pi_{j+1} = \phi_j(\phi_{j+1} + \beta_j \pi_j) = \phi_j\psi_{j+1} + \beta_j \phi_j \pi_j$

3. 把题中给出的多项式等式把 $t \mapsto A$ 并作用在 r_0 上。

(1) 由 $\psi_{j+1}\phi_{j+1} = \psi_j\phi_{j+1} - \eta_j(\psi_{j-1} - \psi_j)\phi_{j+1} - \xi_j t \psi_j \phi_{j+1}$ 得到

$$t_{j+1} = t_j - \eta_j y_j - \xi_j A t_j.$$

(2) 由 $\psi_j\phi_{j+1} = \psi_j\phi_j - \alpha_j t \psi_j \pi_j$ 得到

$$t_j = \psi_j(A)\phi_j(A)r_0 - \alpha_j A p_j.$$

将 $\hat{r}_j := \psi_j(A)\phi_j(A)r_0$ 作为方便的中间量, 则上式可写成

$$t_j = \hat{r}_j - \alpha_j A p_j.$$

(3) 由 $(\psi_{j-1} - \psi_j)\phi_{j+1} = \psi_{j-1}\phi_j - \psi_j\phi_{j+1} - \alpha_j t \psi_{j-1}\pi_j$ 得到

$$y_j = \psi_{j-1}(A)\phi_j(A)r_0 - t_j - \alpha_j A s_j.$$

(4) 由 $\psi_{j+1}\pi_{j+1} = \psi_{j+1}\phi_{j+1} - \beta_j \eta_j \psi_{j-1}\pi_j + \beta_j(1 + \eta_j)\psi_j\pi_j - \beta_j \xi_j t \psi_j \pi_j$ 作用到 r_0 并代入 (1) 中 $\psi_{j+1}\phi_{j+1}r_0 = t_j - \eta_j y_j - \xi_j A t_j$, 得到

$$p_{j+1} = (t_j - \eta_j y_j - \xi_j A t_j) - \beta_j \eta_j s_j + \beta_j(1 + \eta_j)p_j - \beta_j \xi_j A p_j.$$

(5) 由 $\psi_j \pi_{j+1} = \psi_j \phi_{j+1} + \beta_j \psi_j \pi_j$ 作用到 r_0 :

$$s_{j+1} = t_j + \beta_j p_j.$$

4. 我们已知标准的解更新公式

$$x_{j+1} = x_j + \alpha_j p_j. \quad (4.0)$$

若希望只保留最近两个解向量 x_{j-1}, x_j 并直接用它们构造 x_{j+1} , 需要把 p_j 用 x_{j-1}, x_j 以及其它已保存向量表示。

首先回忆第 3 问中得到的 p_j 的递推:

$$p_j = (t_{j-1} - \eta_{j-1} y_{j-1} - \xi_{j-1} A t_{j-1}) - \beta_{j-1} \eta_{j-1} s_{j-1} + \beta_{j-1} (1 + \eta_{j-1}) p_{j-1} - \beta_{j-1} \xi_{j-1} A p_{j-1}. \quad (4.1)$$

该式中出现了先前的搜索方向 p_{j-1} , 以及 $A p_{j-1}$ 。但由迭代定义

$$x_j = x_{j-1} + \alpha_{j-1} p_{j-1},$$

可得

$$p_{j-1} = \frac{x_j - x_{j-1}}{\alpha_{j-1}}. \quad (4.2)$$

将 (4.2) 代入 (4.1) 中含 p_{j-1} 的那一项, 得到 p_j 表示为仅含 $(x_j - x_{j-1})$ 和其它已知向量的式子:

$$p_j = (t_{j-1} - \eta_{j-1} y_{j-1} - \xi_{j-1} A t_{j-1}) - \beta_{j-1} \eta_{j-1} s_{j-1} + \beta_{j-1} (1 + \eta_{j-1}) \frac{x_j - x_{j-1}}{\alpha_{j-1}} - \beta_{j-1} \xi_{j-1} A p_{j-1}. \quad (4.3)$$

现在将 (4.3) 代入解更新 (4.0):

$$\begin{aligned} x_{j+1} &= x_j + \alpha_j p_j \\ &= x_j + \alpha_j \left[t_{j-1} - \eta_{j-1} y_{j-1} - \xi_{j-1} A t_{j-1} - \beta_{j-1} \eta_{j-1} s_{j-1} \right. \\ &\quad \left. + \beta_{j-1} (1 + \eta_{j-1}) \frac{x_j - x_{j-1}}{\alpha_{j-1}} - \beta_{j-1} \xi_{j-1} A p_{j-1} \right]. \end{aligned}$$

将含 $x_j - x_{j-1}$ 的项展开并把关于 x_j, x_{j-1} 的系数集中, 得到:

$$\begin{aligned} x_{j+1} &= \left(1 + \frac{\alpha_j \beta_{j-1} (1 + \eta_{j-1})}{\alpha_{j-1}} \right) x_j - \frac{\alpha_j \beta_{j-1} (1 + \eta_{j-1})}{\alpha_{j-1}} x_{j-1} \\ &\quad + \alpha_j (t_{j-1} - \eta_{j-1} y_{j-1} - \xi_{j-1} A t_{j-1}) - \alpha_j \beta_{j-1} \eta_{j-1} s_{j-1} - \alpha_j \beta_{j-1} \xi_{j-1} A p_{j-1}. \end{aligned} \quad (4.4)$$

上式即为把 x_{j+1} 用 x_j, x_{j-1} 与一组已保存向量 $(t_{j-1}, y_{j-1}, s_{j-1}, A t_{j-1}, A p_{j-1})$ 表示的显式公式。

5. 下面给出 α_j 与 β_j 的可计算表达式.

取固定影子向量 \tilde{r} (任意非零向量, 常取 $\tilde{r} = r_0$ 或 $\tilde{r} =$ 初始残差的复制)。

$$\hat{r}_j := \psi_j(A) \phi_j(A) r_0.$$

将要构造的中间向量 $t_j = \psi_j \phi_{j+1} r_0$ 满足

$$t_j = \hat{r}_j - \alpha_j A p_j.$$

仿照 BiCG 的做法, 为了确定 α_j 我们对上式取与影子向量 \tilde{r} 的内积并要求

$$\langle \tilde{r}, t_j \rangle = 0.$$

由此得

$$0 = \langle \tilde{r}, t_j \rangle = \langle \tilde{r}, \hat{r}_j \rangle - \alpha_j \langle \tilde{r}, A p_j \rangle.$$

因此我们得到显式可计算的公式:

$$\alpha_j = \frac{\langle \tilde{r}, \hat{r}_j \rangle}{\langle \tilde{r}, A p_j \rangle}. \quad (\text{A})$$

上式的分子要求 $\hat{r}_j = \psi_j(A) \phi_j(A) r_0$ 的内积。为避免每步都从多项式直接构造 \hat{r}_j , 我们利用关系

$$t_j = \hat{r}_j - \alpha_j A p_j,$$

对两端取与 \tilde{r} 的内积并用 $\langle \tilde{r}, t_j \rangle = 0$ 得到与 (A) 一致的解。

经典 BiCG 的 β_j 可由下一步左残差对上一残差的比值得到。

$$\beta_j = \frac{\langle \tilde{r}, \hat{r}_{j+1} \rangle}{\langle \tilde{r}, \hat{r}_j \rangle} \cdot \frac{\alpha_j}{\alpha_{j-1}} \quad (\text{B})$$

□

PROBLEM II Prove that the vectors r_j and r_i^* produced by the BCG algorithm are orthogonal to each other when $i \neq j$, while the vectors p_i and p_j^* are A -orthogonal, i.e.,

$$(A p_j, p_i^*) = 0 \quad \text{for } i \neq j.$$

SOLUTION. 令 $A \in \mathbb{R}^{n \times n}$ 。BCG 产生两组向量序列

$$\{r_j\}_{j \geq 0}, \quad \{p_j\}_{j \geq 0} \quad \text{和} \quad \{r_j^*\}_{j \geq 0}, \quad \{p_j^*\}_{j \geq 0},$$

它们满足常见的递推关系:

$$p_j = r_j + \beta_{j-1} p_{j-1}, \quad (6)$$

$$p_j^* = r_j^* + \beta_{j-1} p_{j-1}^*, \quad (7)$$

$$r_{j+1} = r_j - \alpha_j A p_j, \quad (8)$$

$$r_{j+1}^* = r_j^* - \alpha_j A^T p_j^*, \quad (9)$$

其中系数定义为:

$$\alpha_j = \frac{(r_j, r_j^*)}{(A p_j, p_j^*)}, \quad \beta_j = \frac{(r_{j+1}, r_{j+1}^*)}{(r_j, r_j^*)}. \quad (10)$$

要证明对于任意 $i \neq j$ 有

$$(r_j, r_i^*) = 0, \quad \text{且} \quad (Ap_j, p_i^*) = 0.$$

由于 r_j, r_i^*, p_j, p_i^* 是对称的过程, 故只需证明 $\forall i < j, (r_j, r_i^*) = 0, (Ap_j, p_i^*) = 0$

设每个非负整数 $k \in \mathbb{N}$, 令 $P(k)$: 对任意 $i, j \leq k$ 且 $i \neq j$, 有 $(r_j, r_i^*) = 0$ 且 $(Ap_j, p_i^*) = 0$ 。

1. 当 $k = 0$ 时命题平凡成立, 由于不存在不等的索引对。
2. 假设对于某 $n \geq 0$, 命题 $P(n)$ 成立; 即对任意 $i, j \leq n$ 且 $i \neq j$, 有

$$(r_j, r_i^*) = 0, \quad (Ap_j, p_i^*) = 0.$$

我们要证明 $P(n+1)$: 即只需对任意 $m \leq n$ 有

$$(r_{n+1}, r_m^*) = 0, \quad (Ap_{n+1}, p_m^*) = 0.$$

先证明 $(r_{n+1}, r_m^*) = 0$ 对任意 $m \leq n$ 由 (8),

$$r_{n+1} = r_n - \alpha_n Ap_n.$$

对任意固定 $m \leq n$ 取内积得

$$(r_{n+1}, r_m^*) = (r_n, r_m^*) - \alpha_n (Ap_n, r_m^*). \quad (11)$$

若 $m \leq n-1$, 根据归纳假设 $(r_n, r_m^*) = 0$ 。同时

$$(Ap_n, r_m^*) = (p_n, A^T r_m^*).$$

注意由 (9) 对 $j = m$ 有

$$A^T p_m^* = \frac{r_m^* - r_{m+1}^*}{\alpha_m},$$

因此 $A^T r_m^*$ 落在由 $\{p_0^*, \dots, p_m^*\}$ 张成的子空间中, 而归纳假设给出对这些较小索引的 p_n 与 Ap_t^* ($t \leq n-1$) 正交, 推出 $(p_n, A^T r_m^*) = 0$, 从而 $(Ap_n, r_m^*) = 0$ 。所以右端为 0, 从而 $(r_{n+1}, r_m^*) = 0$ 。

若 $m = n$, 直接在 (11) 中代入 $m = n$ 得

$$(r_{n+1}, r_n^*) = (r_n, r_n^*) - \alpha_n (Ap_n, r_n^*).$$

由 (10), 有 $(Ap_n, p_n^*) = (r_n, r_n^*)/\alpha_n$ 。注意根据归纳假设 r_n^* 与 Ap_n 的唯一非零内积就是 (Ap_n, p_n^*) , 因此在 $m = n$ 情形也成立。综上, 对任意 $m \leq n$ 都有 $(r_{n+1}, r_m^*) = 0$ 。

再证明 $(Ap_{n+1}, p_m^*) = 0$ 对任意 $m \leq n$ 由 (6) 和线性性

$$Ap_{n+1} = Ar_{n+1} + \beta_n Ap_n.$$

于是

$$(Ap_{n+1}, p_m^*) = (Ar_{n+1}, p_m^*) + \beta_n (Ap_n, p_m^*). \quad (12)$$

先看第一项:

$$(Ar_{n+1}, p_m^*) = (r_{n+1}, A^T p_m^*).$$

利用 (9) 可得

$$A^T p_m^* = \frac{r_m^* - r_{m+1}^*}{\alpha_m}.$$

因此

$$(r_{n+1}, A^T p_m^*) = \frac{1}{\alpha_m} ((r_{n+1}, r_m^*) - (r_{n+1}, r_{m+1}^*)).$$

但我们已证明对任意 $t \leq n$ 有 $(r_{n+1}, r_t^*) = 0$, 因此第一项为 0。第二项在 $m \leq n-1$ 时由归纳假设 $(Ap_n, p_m^*) = 0$ 而为 0。只剩下 $m = n$ 的情形。将 $m = n$ 代入 (12) 得

$$(Ap_{n+1}, p_n^*) = (r_{n+1}, A^T p_n^*) + \beta_n (Ap_n, p_n^*).$$

那么,

$$(r_{n+1}, A^T p_n^*) = \frac{1}{\alpha_n} ((r_{n+1}, r_n^*) - (r_{n+1}, r_{n+1}^*)).$$

而我们已证明 $(r_{n+1}, r_n^*) = 0$, 所以

$$(r_{n+1}, A^T p_n^*) = -\frac{1}{\alpha_n} (r_{n+1}, r_{n+1}^*).$$

另一方面由 (10),

$$(Ap_n, p_n^*) = \frac{(r_n, r_n^*)}{\alpha_n}.$$

于是

$$\begin{aligned} (Ap_{n+1}, p_n^*) &= -\frac{1}{\alpha_n} (r_{n+1}, r_{n+1}^*) + \beta_n \frac{(r_n, r_n^*)}{\alpha_n} \\ &= \frac{1}{\alpha_n} \left(-(r_{n+1}, r_{n+1}^*) + \beta_n (r_n, r_n^*) \right). \end{aligned}$$

代入 β_n 的定义 $\beta_n = \frac{(r_{n+1}, r_{n+1}^*)}{(r_n, r_n^*)}$, 括号内为零, 从而

$$(Ap_{n+1}, p_n^*) = 0.$$

因此对任意 $m \leq n$ 都有 $(Ap_{n+1}, p_m^*) = 0$ 。

□

PROBLEM III Consider the linear system

$$\begin{bmatrix} A & B \\ B^T & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}, \quad (8.30)$$

in which $c = 0$ and B is of full rank. Define the matrix

$$P = I - B(B^T B)^{-1} B^T.$$

1. Show that P is a projector. Is it an orthogonal projector? What are the range and null spaces of P ?

2. Show that the unknown x can be found by solving the linear system

$$PAPx = Pb, \quad (8.35)$$

in which the coefficient matrix is singular but the system is consistent, i.e., there is a nontrivial solution because the right-hand side is in the range of the matrix (see Chapter 1).

3. What must be done to adapt the Conjugate Gradient Algorithm for solving the above linear system (which is symmetric, but not positive definite)? In which subspace are the iterates generated from the CG algorithm applied to (8.35)?
4. Assume that the QR factorization of the matrix B is computed. Write an algorithm based on the approach of the previous questions for solving the linear system (8.30).

SOLUTION. 1. 记 $P_B := B(B^T B)^{-1} B^T$, 则 $P = I - P_B$. 先证幂等性:

$$P^2 = (I - P_B)^2 = I - 2P_B + P_B^2.$$

但

$$P_B^2 = B(B^T B)^{-1} B^T B(B^T B)^{-1} B^T = B(B^T B)^{-1} (B^T B) (B^T B)^{-1} B^T = P_B,$$

于是 $P^2 = P$, 即 P 为投影矩阵。

再证对称性:

$$P^T = (I - P_B)^T = I - P_B^T = I - P_B = P,$$

因为 $P_B^T = P_B$. 幂等且对称的投影矩阵等价于正交投影矩阵, 故 P 为把向量正交投影到某子空间上的正交投影。

接下来给出其像与核:

- 若 $v \in \text{range}(P)$, 则存在 w 使 $v = Pw$. 由

$$B^T v = B^T Pw = B^T (I - B(B^T B)^{-1} B^T) w = B^T w - B^T w = 0,$$

得 $v \in \text{Null}(B^T)$. 反过来若 $v \in \text{Null}(B^T)$, 则

$$Pv = (I - P_B)v = v - P_B v = v - B(B^T B)^{-1} B^T v = v,$$

故 $v \in \text{range}(P)$. 因此

$$\text{range}(P) = \text{Null}(B^T).$$

- 若 $z \in \text{range}(B)$, 则 $z = B\xi$ 对某向量 ξ , 于是

$$Pz = (I - P_B)B\xi = B\xi - B(B^T B)^{-1} B^T B\xi = 0,$$

即 $\text{range}(B) \subseteq \text{null}(P)$. 反过来若 $Pz = 0$, 则 $z = P_B z = B(B^T B)^{-1} B^T z \in \text{range}(B)$, 因此

$$\text{null}(P) = \text{range}(B).$$

综上, P 为将向量正交投影到 $\text{Null}(B^T)$ 的正交投影矩阵, 且

$$\text{range}(P) = \text{Null}(B^T), \quad \text{null}(P) = \text{range}(B).$$

2. 原系统等价写为:

$$Ax + By = b, \quad B^T x = 0. \quad (13)$$

由约束 $B^T x = 0$ 可知 $x \in \text{Null}(B^T) = \text{range}(P)$, 即 $x = Px$ 。对第一式左右同时左乘 P :

$$P(Ax + By) = Pb.$$

因 B 的列在 $\text{null}(P)$ 上, 故 $PB = 0$, 得

$$PAx = Pb.$$

代入 $x = Px$ 得

$$PAPx = Pb,$$

即约化系统。

奇异性: 因为 $\text{null}(P) = \text{range}(B)$ 非平凡, 故存在非零向量 $v \in \text{range}(B)$ 使得 $Pv = 0$, 于是

$$PAPv = 0,$$

说明 PAP 在整个 \mathbb{R}^n 上一般是奇异的。

一致性: 若原系统 (13) 有解 (x_*, y_*) , 则同样按上面步骤有

$$PAPx_* = Pb.$$

因此该约化系统至少有一个解, 故一致。也就是说, 约化方程的右端 Pb 必位于 PAP 的列空间内, 所以不会出现无解的情况。而如果单独给定任意 P, A, b , 的确不能保证一致, 但这里的 Pb 是从原有解导出的, 因此保证落在列空间内。

3. 共轭梯度法要求系数矩阵为对称正定。在本题中 PAP 在整个空间上可能仅为半正定或奇异, 故不能直接在全空间上对 PAP 施行标准 CG。可以考虑以下策略: 若 A 为正定的,

a) 投影 CG (在约束子空间上): 将迭代限制在 $\mathcal{S} = \text{Null}(B^T) = \text{range}(P)$ 上。若 A 在 \mathcal{S} 上是正定的, 则算子 PAP 在该子空间上是 SPD。首先, PAP 是对称的, 因为

$$(PAP)^T = P^T A^T P^T = PAP.$$

其次, 对任意 $x \in \mathcal{S}$, 由 $Px = x$ 得

$$x^T (PAP)x = x^T P A P x = x^T A x.$$

根据假设, A 在 \mathcal{S} 上正定, 因此

$$x^T (PAP)x = x^T A x > 0, \quad \forall x \in \mathcal{S} \setminus \{0\}.$$

故算子 PAP 在 \mathcal{S} 上对称且正定, 即 PAP 在 \mathcal{S} 上为 SPD。于是可以在此子空间上使用 CG。实现时需要在每次生成残量与搜索方向后应用投影 P , 以保持向量落在 \mathcal{S} 中。

b) 坐标约化 (基底表示): 取 $Z \in \mathbb{R}^{n \times (n-p)}$ 为 $\text{Null}(B^T)$ 的正交基 (例如 QR 里的 Q_2), 写 $x = Zz$, 代入得到尺寸较小的线性系统

$$Z^T A Z z = Z^T b.$$

若 $Z^T A Z$ 为 SPD, 则可对该小系统直接用标准 CG。迭代点实质上属于 $\text{span}(Z) = \text{Null}(B^T)$, 即 CG 迭代生成的向量都在约束子空间内。

当 A 在该子空间上不为正定时，应改用能够处理半正定或不定对称矩阵的方法，如 MIN-RES，或在子空间上做适当正则化。

4. 令对 B 做完整 QR 分解：

$$B = [Q_1 \ Q_2] \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 R,$$

其中 $[Q_1 \ Q_2] \in \mathbb{R}^{n \times n}$ 为正交矩阵， $Q_1 \in \mathbb{R}^{n \times p}$ ， $Q_2 \in \mathbb{R}^{n \times (n-p)}$ ，且

$$\text{range}(B) = \text{span}(Q_1), \quad \text{Null}(B^T) = \text{span}(Q_2).$$

取 $Z = Q_2$ 为 $\text{Null}(B^T)$ 的正交基，代入 $x = Zz$ 得约化系统

$$\tilde{A}z := Z^T A Z z = Z^T b =: \tilde{b}.$$

若 \tilde{A} 为 SPD，则用 CG 求解 $\tilde{A}z = \tilde{b}$ ，恢复

$$x = Zz = Q_2 z.$$

再由原方程第一行回代求 y ：

$$By = b - Ax.$$

因为 $B = Q_1 R$ 且 R 可逆，取最小二乘/精确解：

$$y = R^{-1} Q_1^T (b - Ax).$$

下面给出伪代码，假设能计算 Q_1, Q_2, R ：

步骤 1： 计算 thin QR： $B = Q_1 R$ ，并补出 Q_2 使 $[Q_1 \ Q_2]$ 正交。

步骤 2： 设 $Z = Q_2$ ，构造 $\tilde{A} = Z^T A Z$ ， $\tilde{b} = Z^T b$ 。

步骤 3： 用 CG（或合适的对称解法）求解 $\tilde{A}z = \tilde{b}$ 。

步骤 4： 恢复 $x = Zz$ 。

步骤 5： 计算 $y = R^{-1} Q_1^T (b - Ax)$ 。

为避免在大矩阵上显式构造 \tilde{A} ，可在迭代中用矩阵向量乘法实现隐式操作 $w \mapsto Z^T (A(Zw))$ ，这在 n 很大而 $n - p$ 较小时尤为重要。

□