Итерационные методы для спектральной задачи

Игорь Лобанов

май 2021

Мотивация введения пространств Крылова.

Отношение Рэлея $r(x) = \frac{x \cdot Ax}{x^2}$ оценивает собственные значения:

$$\lambda_1 = \min \operatorname{spec} A = \min_x r(x) \leq \dots \leq \lambda_N = \max \operatorname{spec} A = \max_x r(x).$$

Цепочка разрастающихся пространств уточняет оценки:

$$Q_k = (q_1 | \cdots | q_k), \quad k = 1 \dots N, \quad \operatorname{ran} Q_k \subset \operatorname{ran} Q_{k+1},$$

$$M_k = \max_{\|y\|=1} r(Q_k y), \quad M_1 \leq \dots \leq M_N = \lambda_N,$$

$$m_k = \min_{\|y\|=1} r(Q_k y), \quad \lambda_1 = m_N \leq \dots \leq m_1.$$

Как оптимально выбирать q_{k+1}?

$$\nabla r(x) = \frac{2}{x^2}(Ax - r(x)x) \leftarrow \text{направление наискорейшего роста.}$$

Если $\max_{x \in \operatorname{ran} Q_k} r(x) =: r(u_k)$, то

$$\nabla r(u_k) \in \operatorname{span}(u_k, Au_k) \quad \Rightarrow \quad q_{k+1} := Au_k.$$

Значит пространства Крылова будут хорошим выбором:

$$Q_k = K(A, q_1, k) := \operatorname{span}(q_1, Aq_1, A^2 q_1, \ldots, A^{k-1} q_1).$$

Трехдиагонализация

Если $Q^T A Q = T$ трехдиагональна и $QQ^T = 1_N$, $Q = (q_1|\cdots|q_N)$, то QR факторизация имеет вид:

$$K(A, q_1, N) = QQ^T K(A, q_1, N) = Q(e_1|Te_1|\cdots|T^{N-1}e_1),$$

Значит базис Q можно получить строя трехдиагонализацию, начиная с заданного первого столбца q_1.
Трехдиагонализацию можно получить отражениями Хаусхолдера, но они разрушают разреженную структуру матрицы A.

Алгоритм Ланцоша

$$T = T_N = \begin{pmatrix} \alpha_1 & \beta_1 & & \cdots & 0 \\ \beta_1 & \alpha_2 & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & \beta_{N-1} \\ 0 & \cdots & & \beta_{N-1} & \alpha_N \end{pmatrix}$$

$$AQ = QT \quad \Rightarrow \quad Aq_k = \beta_{k-1} q_{k-1} + \alpha_k q_k + \beta_k q_{k+1}, \quad \beta_0 q_0 := 0.$$

$$Q^T Q = 1 \quad \Rightarrow \quad \alpha_k = q_k^T A q_k = T_{kk}.$$

$$r_k := (A - \alpha_k) q_k - \beta_{k-1} q_{k-1} \quad \Rightarrow \quad q_{k+1} = r_k / \beta_k, \; \beta_k = \pm \| r_k \|_2.$$

Алгоритм Ланцоша прерывается на шаге m, когда $r_m = 0$.

$$m = \operatorname{ran} K(A, q_1, N).$$

Для промежуточных шагов

$$AQ_k = Q_k T_k + r_k e_k^T,$$

$$\operatorname{span} Q_k = K(A, q_1, k), \quad e_{k;j} = \delta_{k,j}.$$

Оценка собственных значений

Для подпространства $S \subset \mathbb{R}^N$ вектор $y \in S$ и $\theta \in \mathbb{R}$ образуют пару Ритца (θ, y) для $A \in \mathrm{Mat}(n \times n, \mathbb{R})$, если $w \cdot (Ay - \theta y) = 0$.

Зафиксируем разложение Шура матрицы T_k:

$$S_k^T T_k S_k = \Theta_k = \mathrm{diag}(\theta_1, \ldots, \theta_k).$$

Обозначим $Y_k = (y_1 | \cdots | y_k) := Q_k S_k$.
Пары Ритца имеют вид (θ_i, y_i):

$$Q_k^T (AY_k - Y_k \Theta_k) = (Q_k^T A Q_k) S_k - Q_K^T (Q_k S_k) \Theta_k = T_k S_k - S_k \Theta_k = 0.$$

Теорема: Минимум $\|AQ_k - Q_k B\|_2$ по всем $B \in \mathrm{Mat}(k \times k, \mathbb{R})$ достигается на $B = T_k = Q_k^T A Q_k$.

Следовательно θ_i дают “лучшую“ оценку собственных чисел, а матрица лучшего приближения оказывается трехдиагональной.

Погрешность приближения собственного вектора

Ранее мы положили $T_k S_k = S_k \Theta_k$, $Y_k = Q_k S_k$.

$$Q_k^T A Q_k = T_k + Q_k^T r_k e_k^T \quad \Rightarrow$$

$$Ay_i - \theta_i y_i = (AQ_k - Q_k T_k) S_k e_i = r_k (e_k^T S_k e_i) \quad \Rightarrow$$

$$\|Ay_i - \theta_i y_i\|_2 = |\beta_k|\,|s_{ki}|.$$

$$S_k^T S_k = 1 \quad \Rightarrow \quad |s_{ki}| \leq 1.$$

Более того

$$\min_{\mu \in \operatorname{spec} A} |\theta_i - \mu| \leq |\beta_k|\,|s_{ki}|.$$

Сходимость

Теорема. Пусть $A \in \mathrm{Mat}(N \times N)$, $A^T = A$ и ее разложение Шура

$$Z^T A Z = \operatorname{diag} \lambda_k, \quad \lambda_1 \geq \cdots \geq \lambda_N, \quad Z = (z_1 | \cdots | z_N).$$

Пусть выполнено k шагов алгоритма Ланцоша и T_k соответствующая трехдиагональная матрица. Если θ_i i-ое собственное значение матрицы T_k, то

$$\lambda_i \geq \theta_i \geq \lambda_i - (\lambda_1 - \lambda_i) \left(\frac{\kappa_i \tan \phi_i}{c_{k-1}(1 + 2\rho_i)} \right)^2,$$

где $c_{k-1}(x)$ – многочлен Чебышева степени $k-1$,

$$\rho_i = \frac{\lambda_i - \lambda_{i+1}}{\lambda_{i+1} - \lambda_n}, \quad \kappa_i = \prod_{j=1}^{i-1} \frac{\theta_j - \lambda_n}{\theta_j - \lambda_i}, \quad \cos \phi_i = |q_1 \cdot z_i.$$

Практические аспекты.

Матрица A никогда не изменяется в алгоритме Ланцоша.
Если A разрежена и имеет ν ненулевых элементов в строке, то один шаг Ланцоша делает $\approx (2\nu + 8)N$ операций с плавающей запятой.
Матрица T_k хранится в паре векторов α, β.
Собственные значения матрицы T_k можно найти QR разложением для симметричных трехдиагональных матриц, поиском нолей $p(\lambda) = \det(T_k - \lambda)$ и т.п.

Ошибки округления

В алгоритме Ланцоша с округлением

$$A\hat{Q}_k = \hat{Q}_k \hat{T}_k + \hat{r}_k e_k^t + E_k,$$

ошибка итераций оценивается

$$\|E_k\|_2 \approx \epsilon \|A\|_2,$$

где ϵ – машинная точность, а крышка обозначает результат операций с плавающей запятой.

Нарушение ортогональности.

$$\hat{\beta}_k \hat{q}_{k+1} \approx \hat{r}_k + w_k,$$

где

$$\|w_k\|_2 \approx \epsilon \|\hat{r}_k\| \approx \epsilon \|A\|_2 \quad \Rightarrow$$

$$|\hat{q}_{k+1} \cdot \hat{q}_i| \approx \frac{|\hat{r}_k \cdot \hat{q}_i| + \epsilon \|A\|_2}{|\hat{\beta}_k\|}.$$

Катастрофическая ошибка, когда $\beta_k \approx 0$.

Поэтому алгоритм долго не использовался, вместо него использовались отражения Хаусхолдера.

Решением может быть полная или частичная ортогонализация векторов q_i.