Грузберг Александр Владимирович, КМБО-01-20

Email: sasha@gruzberg.ru

Telegram: @alexgruzberg

Алгоритм быстрого итеративного уточнения разложения матрицы по сингулярным значениям

Напомним, что разложением матрицы $A \in \mathbb{R}^{m \times n}$ по сингулярным значениям (SVD) называется такое соотношение

$$A = U \Sigma V^T, \tag{1}$$

где $U \in \mathbb{R}^{m \times m}, V \in \mathbb{R}^{n \times n}$ – ортогональные матрицы, составленные из левых сингулярных векторов $u_{(i)} \in \mathbb{R}^m, i = 1, \dots, m$ и правых сингулярных векторов $v_{(i)} \in \mathbb{R}^n, i = 1, \dots, n$ соответственно, а $\Sigma \in \mathbb{R}^{m \times n}$ – прямоугольная диагональная матрица, составленная из сингулярных значений $\sigma_i \in \mathbb{R}$, $i = 1, \dots, n$ (полагается, что $m \geq n$).

Авторы Yuki Uchino, Takeshi Terao и Katsuhisa Ozaki предлагают быстрый (по сравнению с иными описанными в [1]) алгоритм для полного разложения матрицы по сингулярным значениям. Здесь и далее полагается, что сингулярные значения удовлетворяют следующему соотношению:

$$\sigma_1 > \sigma_2 > \sigma_3 > \cdots > \sigma_n,$$

а их приближённые значения $\tilde{\sigma}_i$ таковы, что $\tilde{\sigma}_i \neq \tilde{\sigma}_j$ для $i \neq j$.

На вход алгоритма поступает матрица $A \in \mathbb{R}^{m \times n}$ и матрицы приближённых сингулярных векторов $\hat{U} \in \mathbb{R}^{m \times m}, \hat{V} \in \mathbb{R}^{n \times n}$. На выходе ожидается получение матриц уточнённых сингулярных векторов $\tilde{U} \in \mathbb{R}^{m \times m}, \tilde{V} \in \mathbb{R}^{n \times n}$ и матрицы уточнённых сингулярных значений $\tilde{\Sigma} \in \mathbb{R}^{m \times n}$. Опишем алгоритм пошагово:

1. $P \leftarrow \left(A\hat{V}\right)_h$ //символ $(\cdot)_h$ означает вычисление с двойной точностью;

2. $Q \leftarrow \left(A^T \widehat{U}_1\right)_h$ //$\widehat{U}_1 \in \mathbb{R}^{m \times n}$ – левый блок блочной матрицы $\left(\widehat{U}_1 \widehat{U}_2\right) = \widehat{U}$;

3. $r_{ii} \leftarrow \left(1 - \left(\hat{u}_{(i)}\right)^T \hat{u}_{(i)}\right)_h, for\ (1 \leq i \leq n)$;

4. $s_{ii} \leftarrow \left(1 - \left(\hat{v}_{(i)}\right)^T \hat{v}_{(i)}\right)_h, for\ (1 \leq i \leq n)$;

5. $t_{ii} \leftarrow \left(\left(\hat{u}_{(i)}\right)^T \hat{p}_{(i)}\right)_h, for\ (1 \leq i \leq n)$ // $P = \left(p_{(1)}, p_{(2)}, \dots, p_{(n)}\right)$;

6. $\tilde{\sigma}_i \leftarrow \left(\frac{t_{ii}}{1-(r_{ii}+s_{ii})/2}\right)_h, for\ (1 \leq i \leq n)$;

7. $\tilde{\Sigma}_n \leftarrow diag(\tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n)$;

8. $\tilde{\Sigma} \leftarrow \left(\tilde{\Sigma}_n, O_{n,m-n}\right)^T$ //$O_{n,m-n} \in \mathbb{R}^{n \times (m-n)}$ – матрица нулей;

9. $C_\gamma \leftarrow \left(P - \widehat{U}_1 \tilde{\Sigma}_n\right)_h$;

10. $C_\delta \leftarrow \left(Q - \hat{V} \tilde{\Sigma}_n\right)_h$;

11. $C_\alpha \leftarrow \left(\widehat{U}_1{}^T C_\gamma\right)_l$ //символ $(\cdot)_l$ означает вычисление с одинарной точностью;

12. $C_\beta \leftarrow \left(\hat{V}^T C_\delta\right)_l$;

13. $D \leftarrow \left(\tilde{\Sigma}_n C_\alpha + C_\beta \tilde{\Sigma}_n\right)_h$ // $D = \left(d_{ij}\right)$;

14. $E \leftarrow \left(C_\alpha \tilde{\Sigma}_n + \tilde{\Sigma}_n C_\beta\right)_h$ // $E = \left(e_{ij}\right)$;

15. $\tilde{g}_{ij} \leftarrow \begin{cases} \left(\frac{d_{ij}}{\tilde{\sigma}_j{}^2 - \tilde{\sigma}_i{}^2}\right)_h, (i \neq j) \\ \left(\frac{s_{ii}}{2}\right)_h, (otherwise) \end{cases}, for\ (1 \leq i, j \leq n)$;

16. $\tilde{f}_{ij} \leftarrow \begin{cases} \left(\frac{e_{ij}}{\tilde{\sigma}_j{}^2 - \tilde{\sigma}_i{}^2}\right)_h, (i \neq j) \\ \left(\frac{r_{ii}}{2}\right)_h, (otherwise) \end{cases}, for\ (1 \leq i, j \leq n)$ // $\tilde{F}_{11} = \left(\tilde{f}_{ij}\right)$;

17. $\tilde{F}_{12} \leftarrow \left(\tilde{\Sigma}_n{}^{-1} P^T \widehat{U}_2\right)_h$;

18. $\tilde{F}_{21} \leftarrow \left(\left(\widehat{U}_2{}^T C_\gamma\right)_l \tilde{\Sigma}_n{}^{-1}\right)_h$;

19. $\tilde{F}_{22} \leftarrow \left(\frac{1}{2}\left(I_{m-n} - \widehat{U}_2{}^T \widehat{U}_2\right)\right)_h$ //$I_{m-n} \in \mathbb{R}^{(m-n) \times (m-n)}$ – единичная матрица;

20. $\widetilde{U} \leftarrow \left(\widehat{U} + \left(\widehat{U}\tilde{F}\right)_l\right)_h // \tilde{F} = \begin{pmatrix} \tilde{F}_{11} & \tilde{F}_{12} \\ \tilde{F}_{21} & \tilde{F}_{22} \end{pmatrix}$ – блочная матрица, где $\tilde{F}_{11} \in \mathbb{R}^{n\times n}, \tilde{F}_{12} \in \mathbb{R}^{n\times(m-n)}, \tilde{F}_{21} \in \mathbb{R}^{(m-n)\times n}, \tilde{F}_{22} \in \mathbb{R}^{(m-n)\times(m-n)}$;
21. $\tilde{V} \leftarrow \left(\hat{V} + \left(\hat{V}\tilde{G}\right)_l\right)_h // \tilde{G} = (\tilde{g}_{ij})$.

Общая «стоимость» алгоритма составляет $3m^3 + 2m^2n + 3mn^2 + 4n^3$ операций в лучшем случае и $4m^3 + 4mn^2 + 4n^3$ – в худшем, где

- $2n^3$: $\hat{V}^T C_\delta$;
- $2mn^2$: $A\hat{V}, A^T\widehat{U}_1, \widehat{U}_1^{\ T} C_\gamma$;
- $2mn(m-n)$: $P^T\widehat{U}_2, \widehat{U}_2^{\ T} C_\gamma$;
- $m(m-n)^2$ или $2m(m-n)^2$: $\widehat{U}_2^{\ T}\widehat{U}_2$;
- $2m^3$: $\widehat{U}\tilde{F}$;
- $2n^3$: $\hat{V}\tilde{G}$.

Список литературы

1. Uchino Yuki, Terao Takeshi, Ozaki Katsuhisa. 2022.08.05 – Acceleration of Iterative Refinement for Singular Value Decomposition. 10.21203/rs.3.rs-1931986/v1.