Численные методы
Лекция 6

Е.А. Яревский

09 декабря 2020

Задачи на собственные значения

Задача на собственные значения: найти λ_i и x_i такие, что

$$Ax_i = \lambda_i x_i.$$

(Задача отыскания спектра матрицы). По крайней мере одно собственное значение всегда существует (в C^N).
В зависимости от свойств матрицы A, можно бывает сделать более содержательные утверждения.
Например, если $A = A^\top$ или $A = A^*$, то имеется N собственных значений (с учётом кратности).

Два основных типа задачи на с.з.:

- Частичная задача на с.з. (большие матрицы)
- Полная задача на с.з. (небольшие матрицы)

Трудоёмкость и алгоритмы существенно отличаются для разных типов матриц.

Частичная задача на с.з. Метод прямых итераций

Пусть A – $N \times N$ матрица с N линейно независимыми собственными векторами,

$$Ax_i = \lambda_i x_i, \quad i = 1, \ldots, N,$$

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_N|.$$

Зададим произвольный вектор u^0 и определим последовательность

$$u^k = Au^{k-1} = A^k u^0.$$

Разложим (по линейной независимости):

$$u^0 = \sum_{i=1}^{N} \alpha_i x_i.$$

Тогда

$$u^k = \sum_{i=1}^{N} \alpha_i \lambda_i^k x_i.$$

Предположим далее, что:

1) $\lambda_1 = \lambda_2 = \ldots = \lambda_p$,

2) $|\lambda_1| > |\lambda_{p+i}|$, $i \geq 1$,

3) $\sum_{i=1}^{p} |\alpha_i| \neq 0$.

$$u^k = \lambda_1^k \sum_{i=1}^{p} \alpha_i x_i + \sum_{i=p+1}^{N} \alpha_i \lambda_i^k x_i = \lambda_1^k \left(\sum_{i=1}^{p} \alpha_i x_i + \sum_{i=p+1}^{N} \alpha_i (\frac{\lambda_i}{\lambda_1})^k x_i \right) =$$

$$= \lambda_1^k \left(\sum_{i=1}^{p} \alpha_i x_i + O\left(\left| \frac{\lambda_{p+1}}{\lambda_1} \right|^k \right) \right) = \lambda_1^k \left(v_p + O\left(\left| \frac{\lambda_{p+1}}{\lambda_1} \right|^k \right) \right).$$

Отношение Рэлея

Рассмотрим отношение

$$\frac{(u^{k+1}, u^k)}{(u^k, u^k)}$$

(отношение Рэлея).

Найдём его поведение при больших k:

$$(u^k, u^k) = \lambda_1^{2k}\left((v_p, v_p) + O\left(\left|\frac{\lambda_{p+1}}{\lambda_1}\right|^k\right)\right),$$

$$(u^{k+1}, u^k) = \lambda_1^{2k+1}\left((v_p, v_p) + O\left(\left|\frac{\lambda_{p+1}}{\lambda_1}\right|^k\right)\right).$$

Окончательно находим:

$$\lambda_1^{(k)} \equiv \frac{(u^{k+1}, u^k)}{(u^k, u^k)} = \lambda_1 + O\left(\left|\frac{\lambda_{p+1}}{\lambda_1}\right|^k\right).$$

Собственный вектор можно найти, рассматривая последовательность

$$u^k/||u^k||.$$

Утверждение: если A – симметричная матрица, то

$$\lambda_1^{(k)} = \lambda_1 + O\left(\left|\frac{\lambda_{p+1}}{\lambda_1}\right|^{\mathbf{2k}}\right).$$

Доказательство: вычислим скалярные произведения в отношении Рэлея, пользуясь ортогональностью собственных векторов...

Замечания

- Величина λ_1^k может становиться слишком большой/маленькой. Нужно периодически нормировать u^k.
- Если $|\lambda_1| = |\lambda_2|$, $\lambda_1 \neq \lambda_2$, сходимость отсутствует! Пример:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Поиск нескольких max собственных значений

Как вычислить несколько (немного) max собственных значений?
Пусть вычислены λ_1 и x_1, $||x_1|| = 1$.
Тогда для вычисления λ_2 можно:

- Выбрать u^0: $(u^0, x_1) = 0$. Не очень хорошо: ортогональность ухудшается на каждом шаге итераций!
- Проводить ортогонализацию на каждом шаге итераций:

$$\tilde{u}^k = (I - x_1(x_1, *))u^k.$$

- Построить другую матрицу

$$A^{(2)} = A - \lambda_1 x_1 (x_1, *).$$

После нахождения λ_2 процесс повторяется, ортогональность требуется ко всем уже вычисленным x_i.

Использование сдвигов

Собственные значения матриц A и $f(A)$ связаны как

$$\lambda_i(f(A)) = f(\lambda_i(A)),$$

собственные вектора совпадают.
В частности,

$$\lambda_i(A - \alpha I) = \lambda_i(A) - \alpha,$$
$$\lambda_i((A - \alpha I)^{-1}) = 1/(\lambda_i(A) - \alpha).$$

Можно ли использовать сдвиг $A - \alpha I$ для ускорения или изменения сходимости?
Скорость сходимости определялась отношением $|\lambda_{p+1}/\lambda_1|$.
После применения сдвига:

$$u^k = (\lambda_1 - \alpha)^k \left(\sum_{i=1}^{p} \alpha_i x_i + \sum_{i=p+1}^{N} \alpha_i \left(\frac{\lambda_i - \alpha}{\lambda_1 - \alpha} \right)^k x_i \right).$$

Чтобы скорость сходимости была максимальной, нужно потребовать найти такое α, что

$$\max_{p+1 \leq i \leq N} \left| \frac{\lambda_i - \alpha}{\lambda_1 - \alpha} \right| = \min .$$

Требуется знание о распределении $\{\lambda_i\}$.

При некоторых α, отношения в сумме могут быть >1 – сходимость к другим с.ч.

Например, если $\lambda_i > 0$, возможен поиск минимального с.ч. при выборе сдвига $\alpha > \lambda_1$:

$$0 > \lambda_1 - \alpha \geq \lambda_2 - \alpha \geq \ldots \geq \lambda_N - \alpha.$$

Сходимость в реальных применениях, однако, может быть достаточно медленной.

Метод обратных итераций

Часто нужны минимальные с.ч. или с.ч. в окрестности некоторого числа. Удобно использовать
метод обратных итераций со сдвигом (shift-invert method).

Пусть матрица A – симметричная, для с.ч. выполняются те же предположения, что и при исследовании метода простой итерации.
Пусть w – произвольный вектор, $||w|| = 1$, μ_0 – произвольное число.
Зададим последовательность:

$$(A - \mu_k I)u^{k+1} = w, \quad \mu_{k+1} = (u^{k+1}, Au^{k+1})/(u^{k+1}, u^{k+1}).$$

Разложим w по собственным векторам A:

$$w = \sum_{i=1}^{N} \alpha_i x_i, \quad \sum_{i=1}^{N} |\alpha_i|^2 = 1.$$

$$u^1 = (A - \mu_0 I)^{-1} w = \sum_{i=1}^{N} \frac{\alpha_i}{\lambda_i - \mu_0} x_i.$$

Если μ_0 – хорошее приближение к λ_1, то можно ожидать, что u^1 хорошо приближает с.в. На произвольном шаге:

$$u^{k+1} = (A - \mu_k I)^{-1} w = \sum_{i=1}^{N} \frac{\alpha_i}{\lambda_i - \mu_k} x_i.$$

Отношение Рэлея:

$$\mu_{k+1} = \frac{(u^{k+1}, Au^{k+1})}{(u^{k+1}, u^{k+1})} = \frac{\sum_{i=1}^{N} \frac{\alpha_i^2 \lambda_i}{(\lambda_i - \mu_k)^2}}{\sum_{i=1}^{N} \frac{\alpha_i^2}{(\lambda_i - \mu_k)^2}} = \frac{\frac{\lambda_1}{(\lambda_1 - \mu_k)^2} \sum_{i=1}^{p} \alpha_i^2 + \sum_{i=p+1}^{N} \frac{\alpha_i^2 \lambda_i}{(\lambda_i - \mu_k)^2}}{\sum_{i=1}^{N} \frac{\alpha_i^2}{(\lambda_i - \mu_k)^2}}.$$

Вычисляем $\mu_{k+1} - \lambda_1$ и находим

$$\frac{\mu_{k+1} - \lambda_1}{(\mu_k - \lambda_1)^2} = \frac{\sum_{i=p+1}^{N} \alpha_i^2 \frac{\lambda_i - \lambda_1}{(\lambda_i - \mu_k)^2}}{\sum_{i=1}^{p} \alpha_i^2 + \sum_{i=p+1}^{N} \alpha_i^2 \frac{(\lambda_1 - \mu_k)^2}{(\lambda_i - \mu_k)^2}}.$$

Если $\mu_k \to \lambda_1$ при $k \to \infty$, то

$$\lim_{k\to\infty} \frac{\mu_{k+1} - \lambda_1}{(\mu_k - \lambda_1)^2} = \frac{\sum_{i=p+1}^{N} \frac{\alpha_i^2}{\lambda_i - \lambda_1}}{\sum_{i=1}^{p} \alpha_i^2} = \mathrm{const}.$$

Т.о., если сходимость есть, то сходимость квадратичная.
Если выбрать μ_0 в окрестности λ_j, то при наличии сходимости она вновь будет квадратичной.

Модификации метода обратных итераций

- Можно менять вектор правой части в направлении с.в.

$$(A - \mu_k I)u^{k+1} = \frac{u^k}{||u^k||}.$$

  Доказано, что в этом случае сходимость будет кубической.

- Можно выбрать постоянный сдвиг, $\mu_k \equiv \mu$, и обращать только одну матрицу, $A - \mu I$. Фактически, это метод прямых итераций с матрицей $(A - \mu I)^{-1}$, и он находит с.з., ближайшее к μ. Подходящим выбором μ можно достичь очень быстрой сходимости.

Полная диагонализация

Для вычисления всех собственных значений, можно привести матрицу преобразованием подобия к диагональному виду:

$$A = QDQ^\top,$$

где Q – ортогональная матрица, $Q^\top Q = I$.
С.з. A и D совпадают.
Вычислить Q за конечное число операций невозможно.

Есть две разные стратегии:
1) Строить (бесконечную) последовательность ортогональных преобразований, приводящих матрицу к матрице с очевидным спектром.
2) Привести матрицу (за конечное число преобразований) к более простому виду, а затем искать с.з., экономя на количестве операций.

Метод Якоби

В размерности n рассмотрим матрицу поворота T_{ij}.
Как выбрать последовательность (i_k, j_k) и угол θ_k?
Минимизируем внедиагональные элементы.
Пусть D_A – диагональная матрица с диагональю A. Определим

$$t^2(A) = ||A - D_A||_F^2 = \sum_{i \neq j} |a_{ij}|^2.$$

Потребуем монотонности последовательности: $t^2(A^{(k+1)}) \leq t^2(A^{(k)})$.
Пусть $C = T_{ij}^\top A T_{ij}$, тогда $||C||_F = ||A||_F$.

$$t^2(C) - t^2(A) = ||C - D_C||_F^2 - ||A - D_A||_F^2 =$$

$$= ||T_{ij}^\top A T_{ij} - D_C||_F^2 - ||A - D_A||_F^2 =$$

$$= \left(\sum_{k,l} |a_{kl}|^2 - \sum_{k} |c_{kk}|^2\right) - \left(\sum_{k,l} |a_{kl}|^2 - \sum_{k} |a_{kk}|^2\right) =$$

$$= \sum_{k} (|a_{kk}|^2 - |c_{kk}|^2) = a_{ii}^2 + a_{jj}^2 - c_{ii}^2 - c_{jj}^2.$$

Рассмотрим 2x2 матрицы $\tilde{A}$, $\tilde{C}$.

$$||\tilde{A}||_F^2 = a_{ii}^2 + a_{jj}^2 + 2a_{ij}^2, \quad ||\tilde{C}||_F^2 = c_{ii}^2 + c_{jj}^2 + 2c_{ij}^2.$$

$$t^2(C) - t^2(A) = \left(||\tilde{A}||_F^2 - 2a_{ij}^2\right) - \left(||\tilde{C}||_F^2 - 2c_{ij}^2\right) = 2(c_{ij}^2 - a_{ij}^2).$$

Выразим c_{ij} как результат умножения $T^\top AT$:

$$c_{ij} = a_{ij}(\cos^2\theta - \sin^2\theta) + (a_{jj} - a_{ii})\cos\theta\sin\theta = a_{ij}\cos 2\theta + \frac{a_{jj} - a_{ii}}{2}\sin 2\theta.$$

Стандартный подход – обнулить c_{ij}. Тогда

$$\tan 2\theta = \frac{2a_{ij}}{a_{jj} - a_{ii}} = \tau.$$

$$t = \tan\theta, \quad t = \frac{-1 \pm \sqrt{1+\tau^2}}{\tau}.$$

$$\cos\theta = \frac{1}{\sqrt{1+t^2}}, \quad \sin\theta = \frac{t}{\sqrt{1+t^2}}.$$

Стратегии перебора индексов

1) Исключить максимальное по модулю значение.
Пусть I, J: $|a_{IJ}| \geqslant \max_{i \neq j} |a_{ij}|$. Тогда

$$|a_{IJ}| \geqslant \frac{1}{n(n-1)} t^2(A^{(k)}).$$

$$t^2(A^{(k+1)}) = t^2(A^{(k)}) - 2a_{IJ}^2 \leqslant t^2(A^{(k)}) \left(1 - \frac{2}{n(n-1)}\right) \leqslant$$

$$\leqslant t^2(A^{(0)}) \left(1 - \frac{2}{n(n-1)}\right)^{k+1}.$$

Фактически, сходимость оказывается асимптотически квадратичной:

$$t^2(A^{(k+1)}) \leqslant C \left[t^2(A^{(k)})\right]^2.$$

2) Последовательный перебор индексов ($n(n-1)/2$ вращений).
3) Пороговый перебор (одновременно выбираются несколько максимальных по модулю элементов).

QR-алгоритм

Запишем QR-разложение в виде:

$$A = Q_1 R_1,$$

где Q_1 – ортогональная, R_1 – правая треугольная матрицы.
$R_1 = Q_1^{-1} A$, поэтому $A_1 = R_1 Q_1 = Q_1^{-1} A Q_1$ подобна A.
Построим последовательность матриц:

$$A_n = Q_{n+1} R_{n+1}, \quad A_{n+1} = R_{n+1} Q_{n+1} = Q_{n+1}^{-1} A_n Q_{n+1}.$$

Все матрицы A_{n+1} подобны, и подобны A.

Сходимость QR-алгоритма: иллюстрация основной идеи

Сходимость QR-алгоритма

Теорема. *Пусть все диагональные миноры матрицы A не вырождены. Тогда последовательность матриц A_n при $n \to \infty$ сходится к клеточному правому треугольному виду, где клетки с одинаковыми по модулю с.зн. имеют размер, совпадающий с кратностью с.зн.*

Указанный вид матрица приобретает после, возможно, подходящей перестановки строк и столбцов.

Если все собственные значения различны по модулю, матрицы A_n стремятся к диагональной матрице с с.зн. на диагонали.

QR-итерации со сдвигом

Выберем сдвиг σ_n.
Построим QR-разложение сдвинутой матрицы:

$$Q_{n+1}R_{n+1} = A_n - \sigma_n I, \quad A_{n+1} = R_{n+1}Q_{n+1} + \sigma_n I.$$

Можно записать в виде:

$$A_{n+1} = Q_{n+1}^\top (A_n - \sigma_n I) Q_{n+1} + \sigma_n I.$$

Свойства Хессенберговости и трёхдиагональности не нарушаются.
Одинарный сдвиг: в качестве σ_n можно взять элемент (N, N) матрицы A_n.
Варианты: сдвиг Уилкинсона (по матрице 2×2).
Для несимметричных матриц – двойной сдвиг (неявный).

Скорость сходимости и сложность

При применении стратегии одинарного сдвига:
для несимметричных матриц – сходимость с.зн. квадратичная, для симметричных матриц – сходимость с.зн. кубическая.

При применении стратегии двойного сдвига:
для несимметричных матриц – сходимость с.зн. кубическая.

Количество операций на одном шаге QR-алгоритма:

- для произвольной матрицы: $O(N^3)$.
- для Хессенберговой матрицы: $O(N^2)$.
- для трёхдиагональной матрицы: $O(N)$.
  Если необходимы собственные векторы, то нужно накапливать ортогональные матрицы, тогда $O(N^2)$ операций.

Приведение матрицы к специальной форме – $O(N^3)$ операций.

Модификации и разновидности

Обобщённая задача на собственные значения:

$$A\mathbf{y} = \lambda B\mathbf{y}.$$

QZ-алгоритм, обобщение QR алгоритма (1973).

LR-алгоритм:

$$A_n = L_{n+1} R_{n+1}, \quad A_{n+1} = R_{n+1} L_{n+1} = L_{n+1}^{-1} A_n L_{n+1}.$$

L – левая треугольная, R – правая треугольная матрицы.

Метод Ланцоша

Рассмотрим большую разреженную симметричную матрицу A. Требуется трёхдиагонализация: $T = Q^\top A Q$.
Возможно использовать метод Хаусхолдера, но не сохраняет разреженность.
Попробуем найти элементы T непосредственно. Пусть $Q = [q_1, q_2, \dots, q_N]$,

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & \dots & 0 \\ \beta_1 & \alpha_2 & \beta_2 & \dots & 0 \\ 0 & \beta_2 & \alpha_3 & \dots & \dots \\ \dots & \dots & \dots & \dots & \beta_{N-1} \\ 0 & 0 & \dots & \beta_{N-1} & \alpha_N \end{pmatrix}.$$

Приравняем столбцы в равенстве $AQ = QT$:

$$Aq_j = \beta_{j-1} q_{j-1} + \alpha_j q_j + \beta_j q_{j+1}, \quad j = 1, \ldots, N-1, \quad \beta_0 q_0 = 0.$$

Из-за ортонормированности векторов q_j: $\alpha_j = q_j^\top A q_j$.
Определим $r_j = (A - \alpha_j I) q_j - \beta_{j-1} q_{j-1}$.
Если r_j ненулевой, то $q_{j+1} = r_j/\beta_j$, $\beta_j = \pm||r_j||_2$.

Итерации Ланцоша:

$$r_0 = q_1; \;\; \beta_0 = 1; \;\; q_0 = 0; \;\; j = 0$$

$$q_{j+1} = r_j/\beta_j; \;\; j = j+1; \;\; \alpha_j = q_j^\top A q_j;$$
$$r_j = (A - \alpha_j I) q_j - \beta_{j-1} q_{j-1}; \;\; \beta_j = ||r_j||_2.$$

q_j – векторы Ланцоша.
Ортогонализация в пространстве Крылова $\{A^k q_1\}$.

Проверка попарной ортогональности

Перепишем основное равенство:

$$\tilde{q}_{j+1} \equiv \beta_j q_{j+1} = (A - \alpha_j) q_j - \beta_{j-1} q_{j-1}.$$

1) Определение α_j

$$(\tilde{q}_{j+1}, q_j) = ((A - \alpha_j) q_j - \beta_{j-1} q_{j-1}, q_j) = (Aq_j, q_j) - \alpha_j = 0.$$

2) Определение β_{j-1}

$$(\tilde{q}_{j+1}, q_{j-1}) = ((A - \alpha_j) q_j - \beta_{j-1} q_{j-1}, q_{j-1}) = (Aq_j, q_{j-1}) - \beta_{j-1} = 0.$$

3) $k \leq j - 2$:

$$(\tilde{q}_{j+1}, q_k) = ((A - \alpha_j) q_j - \beta_{j-1} q_{j-1}, q_k) = (Aq_j, q_k) = (q_j, Aq_k) =$$

$$= (q_j, \beta_{k-1} q_{k-1} + \alpha_k q_k + \beta_k q_{k+1}) = 0 + 0 + 0 = 0.$$

Оценки Каниэля-Пейджа

Теорема. *Пусть A – симметричная $N \times N$ матрица с с.зн. $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ и соответствующими с.вект. z_i. Если $\theta_1 \geq \ldots \geq \theta_j$ есть с.зн. матрицы T_j, полученной после j шагов итераций Ланцоша, то*

$$\lambda_1 \geq \theta_1 \geq \lambda_1 - \frac{(\lambda_1 - \lambda_N)\tan^2(\varphi_1)}{C_{j-1}^2(1 + 2\rho_1)},$$

где $\cos(\varphi_1) = |q_1^\top z_1|$, $\rho_1 = (\lambda_1 - \lambda_2)/(\lambda_2 - \lambda_N)$, C_{j-1} *– полином Чебышева степени* $j - 1$.

Аналогичная оценка получена для λ_N, и худшие – для внутренних с.зн.

Проблема алгоритма Ланцоша: потеря ортогональности.

SVD разложение

Обратимся к исследованию прямоугольных матриц.
Метод наименьших квадратов (Least Square), LS с ограничениями, метод главных компонент (Principal Component Analysis, PCA).

Важное средство работы с такими матрицами (а также и с квадратными) – сингулярное разложение матрицы (Singular Value Decomposition, SVD).

Для матрицы A, $m \times n$, $m \geq n$, сингулярным разложением называется представление

$$A = U \, \Sigma \, V^\top.$$

Здесь

$$U^\top U = UU^\top = I_{m\times m}, \quad V^\top V = VV^\top = I_{n\times n}, \quad \Sigma = \left[\begin{array}{ccc} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \\ \hline & 0 & \end{array}\right]_{m\times n}.$$

где вещественные значения σ_i: $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$.
Ранг матрицы A равен числу ненулевых значений.

$$
\begin{aligned}
AA^\top &= U\Sigma V^\top V\Sigma^\top U^\top = U\Sigma\Sigma^\top U^\top \\
&= U \left[\begin{array}{ccc|c} \sigma_1^2 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_n^2 & \\ \hline & 0 & & 0 \end{array}\right]_{m\times m} U^\top .
\end{aligned}
$$

Столбцы U (левые сингулярные вектора) – собственные вектора $AA^\top$ с собственными значениями σ_i^2 и 0.

$$
A^\top A = V\,\Sigma^\top\,\Sigma V^\top .
$$

Аналогично, столбцы V (правые сингулярные вектора) – собственные вектора $A^\top A$ с собственными значениями σ_i^2.

Построение матрицы меньшего ранга

$$A = U\,\Sigma\,V^* = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^*.$$

(Сумма r штук матриц ранга 1.)
Найдем такую матрицу B, что $\mathrm{rank}(B) = k < r$, и

$$||A - B||_F = \min.$$

Решение этой задачи:

$$B = A_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^*.$$

$$||A - B||_F = \sigma_{k+1}^2 + \ldots + \sigma_r^2.$$

Построение матрицы меньшего ранга

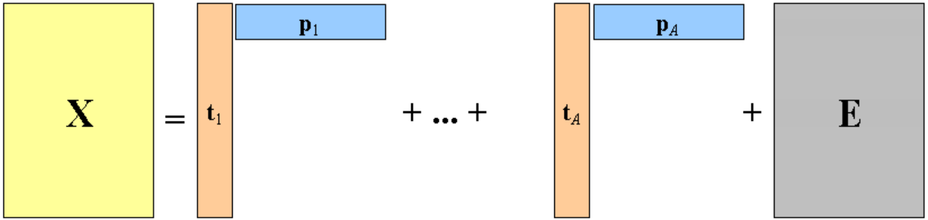

Рис. 5 Разложение по главным компонентам

Похожая задача:
пусть $A = \hat{A} + E$, где E – ошибка в матрице A, и известна оценка $||E|| < \varepsilon$.
Тогда требуется найти матрицу B минимального ранга такую, что

$$||A - B||_F < \varepsilon.$$

Псевдообратная матрица

SVD может быть использовано для вычисления псевдообратной матрицы A^+:

$$A^+ = V\Sigma^+ U^\top.$$

Псевдообратная Σ^+ формируется заменой всех ненулевых диагональных элементов их обратными.

Если A несингулярна, псевдообратная совпадает с обычной обратной.

Построение псевдообратной матрицы — способ решения задачи наименьших квадратов.

Метод наименьших квадратов

Для матрицы A, $m \times n$, с рангом r, и вектора длины m, найти вектор x минимальной длины, минимизирующий $||b - Ax||_2$.

$$||b - Ax||_2 = ||b - U\,\Sigma\,V^\top x||_2 = ||U^\top b - \Sigma\,y||_2.$$

Обозначим

$$U^T\mathbf{b} = \begin{bmatrix} c_1 \\ \vdots \\ c_r \\ d_1 \\ \vdots \\ d_{m-r} \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix},$$

тогда

Метод наименьших квадратов

$$||b - Ax||_2^2 = \sum_{i=1}^{r} (c_i - \sigma_i y_i)^2 + \sum_{i=r+1}^{n} d_i^2.$$

Минимум достигается при $y_i = c_i/\sigma_i$.
$x = Vy$, $||x||_2 = ||y||_2$, так что минимальная длина достигается при $y_i = 0$ для $i = r+1, \ldots, n$.

Вычисление сингулярного разложения

1) Наиболее очевидный способ: решение спектральной задачи для $AA^\top$, вычисление U и V.
   Проблема: спектральное разложение более чувствительно к возмущениям! (число обусловленности возводится в квадрат).
2) Правильный способ: решение спектральной задачи

$$H = \begin{bmatrix} 0 & A^\top \\ A & 0 \end{bmatrix}.$$

   Число обусловленности не увеличивается, но размер матрицы удваивается.
3) Наиболее эффективный способ: Двухдиагонализация (Голуб, Кахан, 1965), далее – неявный симметричный QR-алгоритм для матрицы $A^\top A$.

Упражнение 4: Оценки собственных значений матрицы

Рассмотрим матрицу A, сформированную в упражнении 3.
Для этой матрицы:
1) Найдите максимальное собственное значение с помощью метода прямых итераций. Постройте его зависимость от N, определите тип его асимптотического поведения при больших N (степень).
2) Найдите минимальное собственное значение с помощью метода прямых итераций со сдвигом. Постройте зависимость спектрального числа обусловленности от N.
3) Определите скорости сходимости методов Якоби и Зейделя для A как функции N, воспользовавшись матричными формулировками методов. Постройте зависимости $1 - q$ от N в подходящем масштабе, сравните скорости методов.

Литература

1. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М., *Численные методы*, М.: Бином. Лаборатория знаний, 2012. - 636 с. Параграфы 6.6, 6.7, 6.11, 6.13.

2. Голуб Д.Х, Ван Лоун Ч.Ф., *Матричные вычисления*, М.: Мир, 1999. - 548 с. Параграфы 5.4, 8.2, 8.3, 9.1.