МАГОЛЕГО
Линейная алгебра в приложениях
Семинар 11

Д. И. Пионтковский, И. К. Козлов
(*ВШЭ*)

9 апреля 2021

Литература

Шевцов Г.С.,
Линейная алгебра: теория и прикладные аспекты,
Глава 10. Итерационные методы,

Метод итераций

Семинар 11

Решение уравнения

$$x = Px + c$$

методом итераций:

$$x^{(n+1)} = Px^{(n)} + c.$$

Семинар 11

Утв 1. Процесс сходится при любом начальном $x^{(0)}$ $\Leftrightarrow$ спектральный радиус $\rho(P) < 1$.

Утв. 2 Верна оценка:

$$\left|x - x^{(k)}\right| \leq \frac{\|P\|^k}{1 - \|P\|} |x^{(1)} - x^{(0)}|.$$

Если $\|P\| < 1$, то метод итераций сходится.

Матричную норму можно брать любой.

Семинар 11

Применим оценку

$$\left|x - x^{(k)}\right| \leq \frac{\|P\|^k}{1 - \|P\|}\left|x^{(1)} - x^{(0)}\right|.$$

- Если $x^0 = 0$, то $x^{(1)} = c$ и

$$\left|x - x^{(k)}\right| \leq \frac{\|P\|^k}{1 - \|P\|}|c|.$$

- Если $x^{(0)} = c$, то $x^{(1)} = Pc + c$ и

$$\left|x - x^{(k)}\right| \leq \frac{\|P\|^{k+1}}{1 - \|P\|}\left|x^{(0)}\right|.$$

Семинар 11

Задача. Методом итераций решить систему

$$\begin{cases} 10x_1 + 2x_2 + x_3 = 10, \\ x_1 + 10x_2 + 2x_3 = 12, \\ x_1 + x_2 + 10x_3 = 8. \end{cases}$$

Решение. Переходим к системе:

$$\begin{cases} x_1 = 1 - 0,2x_2 - 0,1x_3, \\ x_2 = 1,2 - 0,1x_1 - 0,2x_3, \\ x_3 = 0,8 - 0,1x_1 - 0,1x_2. \end{cases}$$

Семинар 11

Матрица

$$P = \begin{pmatrix} 0 & -0,2 & -0,1 \\ -0,1 & 0 & -0,2 \\ -0,1 & -0,1 & 0 \end{pmatrix}.$$

Заметим, что

$$\|P\|_1 = \max(0,2; \quad 0,3; \quad 0,3) = 0,3 < 1,$$

поэтому метод итераций сойдётся.

Норму $\|P\|_1$ легко считать (сумма модулей по столбцам).

За начальное приближение можно взять $x^{(0)} = c$:

$$x_1^{(0)} = 1, \qquad x_2^{(0)} = 1,2, \qquad x_3^{(0)} = 0,8.$$

Семинар 11

Решение

$$x \approx x^{(5)} \approx (0,737; \quad 1,001; \quad 0,626)$$

Оценка погрешности

$$\left| x - x^{(5)} \right| \leq \frac{\|P\|^6}{1 - \|P\|} \left| x^{(0)} \right| = \frac{(0,3)^6}{0,7} 3 \leq 0,001.$$

Приведение к виду, удобному для итераций

1. Метод итераций.
2. Приведение к виду, удобному для итераций.
3. Метод Зейделя.
4. Добавления. Другие примеры методы итераций.

Семинар 11

Задача. Привести к виду, удобному для итераций

$$Ax = b.$$

Случай-1. Пусть $A = A^T$ и есть диагональное преобладание:

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

Диагональные элементы больше суммы модулей остальных элементов в своей строке.

Выражаем из i-того уровнения x_i:

$$\begin{cases} x_1 = \ldots, \\ x_2 = \ldots, \\ \ldots \end{cases}$$

Семинар 11

Задача. Привести к виду, удобному для итераций

$$Ax = b.$$

Случай-2. Пусть $A = A^T$ и все собственные числа $\lambda_i > 0$.
Берём

$$P = E - \tau A, \qquad \bar{b} = \tau b,$$

где

$$\tau = \frac{2}{\lambda_{\min} + \lambda_{\max}}.$$

Затем решаем методом итераций

$$x_{k+1} = Px_k + \bar{b}.$$

Случай-3. Если $A \neq A^T$, то переходим к системе

$$A^T A x = A^T b.$$

Далее поступаем так же, как и для симметричной матрицы.

Семинар 11

На практике находить

$$\tau = \frac{2}{\lambda_{\min} + \lambda_{\max}}$$

иногда затруднительно. Поэтому берут

$$\tau = \frac{2}{\sigma},$$

где σ больше какой-нибудь из просто вычислимых матричных норм

$$\|A\|_1 = \max_j \sum_i |a_{ij}|, \qquad \|A\|_\infty = \max_i \sum_j |a_{ij}|.$$

Семинар 11

Метод Зейделя.

Пусть мы решаем уравнение $Ax = y$. Разложим

$$(\mathrm{L} + \mathrm{D})\vec{x} = -\mathrm{U}\,\vec{x} + \vec{b},$$

где D — диагонаяльная, а L и U — строго нижнетреугольная и верхнетреугольная матрицы. Итерационный процесс в методе Гаусса — Зейделя:

$$(\mathrm{L} + \mathrm{D})\vec{x}^{(k+1)} = -\mathrm{U}\,\vec{x}^{(k)} + \vec{b}.$$

Семинар 11

По сути это тот же метод итераций:

$$(\mathrm{L}+\mathrm{D})\vec{x}^{k+1} = -\mathrm{U}\,\vec{x}^{k} + \vec{b}$$

эквивалентная

$$\vec{x}^{k+1} = (\mathrm{L}+\mathrm{D})^{-1}\vec{b} - (\mathrm{L}+\mathrm{D})^{-1}\mathrm{U}\,\vec{x}^{k}.$$

Так что метод сходится, если

$$\left\|(\mathrm{L}+\mathrm{D})^{-1}\,\mathrm{U}\right\| < 1.$$

Другие примеры методы итераций

1. Метод итераций.
2. Приведение к виду, удобному для итераций.
3. Метод Зейделя.
4. Добавления. Другие примеры методы итераций.

Семинар 11

Пусть $A = A^T$.

Методом итераций можно находить собственные значения A.

Возьмём базис e_i из собственных векторов

$$Ae_i = \lambda_i e_i, \qquad |\lambda_1| \geq |\lambda_2| \geq \ldots.$$

Тогда

$$A^k(e_1, \ldots, e_n) = \left(\lambda_1^k e_1, \ldots, \lambda_n^k e_n\right) = \lambda_1^k \cdot \left(e_1, \left(\frac{\lambda_2}{\lambda_1}\right)^k e_2, \ldots, , \left(\frac{\lambda_n}{\lambda_1}\right)^k e_n\right).$$

$\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0$, если $\left|\frac{\lambda_j}{\lambda_1}\right| < 1$.

Семинар 11

Берём

$$v_{k+1} = \frac{Av_k}{\|Av_k\|}.$$

В пределе получаем собственный вектор

$$Av = \lambda v,$$

после этого переходим к ортогональному дополнению

$$(w, v) = 0.$$

Семинар 11

Утверждение. Градиент — направление наибольшего роста функции.

Дана функция $f(x^1, \ldots, x^n) : \mathbb{R}^n \to \mathbb{R}$. Её градиент:

$$\left(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}\right).$$

Производная по направлению v (в нуле):

$$\frac{d}{dt} f(tv^1, \ldots, tv^n) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} v^i = (\operatorname{grad} f, v).$$

Напомним, что скалярное произведение — произведение длин на косинус угла между векторами:

$$(u, v) = |u||v| \cos\varphi$$

Наибольшее скалярное произведение — если векторы сонаправлены.

Семинар 11

Линейная регрессия.

Даны векторы $x_i \in \mathbb{R}^n$ и предсказываемые значения y_i. Пусть мы хотим максимально точно приблизить

$$y_i = (x_i, \omega) \, .$$

Вариант-1. Точное решение (через псевдообратную матрицу)

$$\omega = X^{+} y.$$

Семинар 11

Вариант-2. Градиентный спуск.

Мы минимизируем функцию потерь:

$$L(\omega, X, y) = \frac{1}{n}\sum_i (y_i - (x_i, \omega))^2.$$

Метод итераций:

$$\omega \to \omega - \alpha \frac{\partial}{\partial \omega} L.$$

Аналогично для логистической регрессии.

Семинар 11

Вариант-3. Максимизация правдоподобия.

$$y_i = (x_i, \omega) + \varepsilon,$$

где шум ε нормально распределён $N(0, \sigma^2)$, т.е.

$$p(y|x,\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-(x,\omega))^2}{2\sigma^2}\right).$$

Семинар 11

Максимизация функции правдоподобия

$$Likelihood(y, X, \omega) = \prod_i p(y_i | x_i, \omega) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - (x_i, \omega)^2)}{2\sigma^2} \right)$$

эквивалентна методу наименьших квадратов

$$\min_{\omega} (y_i - (x_i, \omega)^2).$$

Это легко увидеть, взяв логарифм функции правдоподобия.

Вероятностный подход/Байесовские методы.