Математическое программирование.
Лекция 2. Решение СЛАУ

Матрично-векторный формализм

Рассмотрим для примера простую систему линейных уравнений, в которой три переменных и три уравнения:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = c_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = c_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = c_3 \end{cases} .$$

Матрично-векторное представление этой системы линейных алгебраических уравнений может быть записано как

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} * \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

или еще проще,

$$\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{c} \quad .$$

В последнем равенстве, предполагается, что векторы переменных и свободных членов уравнений представлены в виде матриц-столбцов. Однако, ничего не мешает использовать и формализм матриц-строк для представления векторов. В этом случае систему можно записать в виде

$$\boldsymbol{x}^{\boldsymbol{T}}\ \boldsymbol{A}^{\boldsymbol{T}} = \boldsymbol{c}^{\boldsymbol{T}} \quad ,$$

где $\boldsymbol{x}^{\boldsymbol{T}}=(x_1, x_2, x_3)$, $\boldsymbol{c}^{\boldsymbol{T}}=(c_1, c_2, c_3)$.
В общем случае, размерность может быть любой, при соблюдении правил умножения матриц, число столбцов в матрице $\boldsymbol{A}$ должно быть равно числу строк в матрице $\boldsymbol{x}$. Матрица $\boldsymbol{A}$ в общем случае необязательно квадратная.

Погрешности

Пусть u и $\hat{u}$ - точное и приближенное значение некоторой величины. Тогда *абсолютной погрешностью* приближения $\hat{u}$ называют величину Δu , которая удовлетворяет неравенству

$$|u-\hat{u}| \le \Delta u \quad .$$

Относительной погрешностью называют величину δu , которая удовлетворяет неравенству

$$\left|\frac{u-\hat{u}}{u}\right| \le \delta u \quad .$$

Оценим для примера погрешность Δx в определении корня $x=1$ уравнения

$$a x^4 + b x^3 + c x + d = 0 \quad ,$$

с коэффициентами

$$a = b = 1 \pm 10^{-3}, \quad c = d = -1 \pm 10^{-3}$$

не решая само уравнение. Погрешность параметров дает погрешность корней, поэтому уравнение можно представить в виде

$$(a+\Delta a)(x+\Delta x)^4 + (b+\Delta b)(x+\Delta x)^3 + (c+\Delta c)(x+\Delta x) + (d+\Delta d) = 0 \quad ,$$

а с другой стороны имеет место точное равенство

$$a x^4 + b x^3 + c x + d = 0 \quad ,$$

поэтому вычитая последнее из предпоследнего уравнения и пренебрегая величинами с порядком малости больше единицы, можно получить следующую оценку:

$$\Delta x \approx \left| \frac{-\Delta a x^4 - \Delta b x^3 - \Delta c x - \Delta d}{4 a x^3 + 3 b x^2 + c} \right|$$

Если сюда подставить значения параметров, их погрешности и значение корня, можно получить, что для корня $x=1$, $\Delta x = \frac{2}{3} 10^{-3}$. Этот пример достаточно наивен, в реальных задачах приходится учитывать вероятностные распределения для значений параметров, изучать распространение ошибки через алгоритм и получать оценку распределения для получаемой в результате величины.

Линейные пространства, нормы векторов и матриц

Множество $\boldsymbol{R^n}$ называется линейным пространством если выполняется три группы аксиом:

I
1. определена операция суммирования, $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{z}, \quad \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{R}$;
2. определено умножение на число, $\lambda \boldsymbol{x} \in \boldsymbol{R^n}, \ \lambda \in \boldsymbol{R}, \ \boldsymbol{x} \in \boldsymbol{R^n}$;
3. существует нулевой элемент, $\boldsymbol{x} + \boldsymbol{0} = \boldsymbol{x}$;
4. для каждого элемента существует обратный, $\boldsymbol{x} + (-\boldsymbol{x}) = \boldsymbol{0}, \quad \boldsymbol{x}, (-\boldsymbol{x}) \in \boldsymbol{R^n}$.

II
1. $1 \cdot \boldsymbol{x} = \boldsymbol{x}$;
2. $\alpha(\beta \boldsymbol{x}) = \alpha\beta(\boldsymbol{x}), \quad \alpha, \beta \in R$.

III
1. $(\alpha+\beta)\boldsymbol{x} = \alpha \boldsymbol{x} + \beta \boldsymbol{x}$;
2. $\alpha(\boldsymbol{x} + \boldsymbol{y}) = \alpha \boldsymbol{x} + \beta \boldsymbol{y}$.

Не важно, как именно выполняются действия, главное, чтобы выполнялись аксиомы.

Набор векторов $\{\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_n}\} \in \boldsymbol{R^n}$ является *линейно независимым*, если из равенства $\sum_{i=1}^{n} \alpha_i \boldsymbol{a_i} = 0$ следует, что все коэффициенты равны нулю, т. е. $\alpha_i = 0 \ \forall \ i=1,2,\ldots,n$.

С каждой матрицей А, размера $m \times n$ связано два важных подпространства. *Область значений*

$$range(\boldsymbol{A}) = \{\boldsymbol{y} \in \boldsymbol{R^m} : \boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x} \ \ \text{для} \ \ \boldsymbol{x} \in \boldsymbol{R^n}\} \quad .$$

Ядро или *нуль-пространство* матрицы:

$$null(\boldsymbol{A}) = \{\boldsymbol{x}\in\boldsymbol{R}^{\boldsymbol{n}} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\} \quad .$$

Если заданы векторы $\{\boldsymbol{a}_{\boldsymbol{1}}, \boldsymbol{a}_{\boldsymbol{2}}, \dots, \boldsymbol{a}_{\boldsymbol{n}}\}$, то множество всех из линейных комбинаций называется *линейной оболочкой,*

$$span\{\boldsymbol{a}_{\boldsymbol{1}}, \boldsymbol{a}_{\boldsymbol{2}}, \dots, \boldsymbol{a}_{\boldsymbol{n}}\} = \left\{\sum_{j=1}^{n} \beta_j \boldsymbol{a}_{\boldsymbol{j}} : \beta_j \in R, j=1,2,..,n\right\} \quad .$$

Ранг матрицы это размерность области значений:

$$rank(\boldsymbol{A}) = dim(range(\boldsymbol{A})) \quad .$$

Ранг матрицы равен максимальному числу линейно независимых строк или столбцов матрицы. Также имеют место равенства

$$rank(\boldsymbol{A}) = rank(\boldsymbol{A}^{\boldsymbol{T}}) \quad ,$$

$$\forall \boldsymbol{A}\in\boldsymbol{R}^{\boldsymbol{m*n}} \quad dim(null(\boldsymbol{A})) + rank(\boldsymbol{A}) = n \quad .$$

Единичная матрица:

$$\boldsymbol{I}_{\boldsymbol{n}} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 \end{pmatrix} = [\boldsymbol{e}_{\boldsymbol{1}}, \boldsymbol{e}_{\boldsymbol{2}}, \dots, \boldsymbol{e}_{\boldsymbol{n}}] \quad ,$$

здесь справа столбцово-векторное представления, составленное из канонических векторов (или базисных векторов).

Если $\boldsymbol{A}, \boldsymbol{X}\in\boldsymbol{R}^{\boldsymbol{n*n}} : \boldsymbol{A}\boldsymbol{X} = \boldsymbol{I}_{\boldsymbol{n}}$, то $\boldsymbol{X}$ - обратная матрица к $\boldsymbol{A}$, обозначается как $\boldsymbol{A}^{-\boldsymbol{1}}$. Если такая матрица существует, то это означает что матрица $\boldsymbol{A}$ *невырожденная*.

Для матрицы $\boldsymbol{A}\in\boldsymbol{R}^{\boldsymbol{n*n}}$ определен детерминант (число)

$$det(\boldsymbol{A}) = a, \quad a\in\boldsymbol{R} \quad .$$

Детерминант определяется рекурсивно через детерминанты матриц размерностью на единицу меньше:

$$det(\boldsymbol{A}) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} det(\boldsymbol{A}_{\boldsymbol{1j}}) \quad ,$$

где $\boldsymbol{A}_{\boldsymbol{1j}}$ - *алгебраическое дополнение* элемента a_{1j} , т. е. матрица полученная из матрицы $\boldsymbol{A}$ путем вычеркивания 1-ой строки и *j*-го столбца. Ниже приведена сводка некоторых полезных свойств детерминанта:

$$det(\boldsymbol{A}\boldsymbol{B}) = det(\boldsymbol{A}) det(\boldsymbol{B}), \quad \boldsymbol{A}, \boldsymbol{B}\in\boldsymbol{R}^{\boldsymbol{n*n}} \quad ,$$

$$det\left(\boldsymbol{A}^{\boldsymbol{T}}\right) \;=\; det\left(\boldsymbol{A}\right) \quad ,$$

$$det\left(c\,\boldsymbol{A}\right) \;=\; c^{n}\,det\left(\boldsymbol{A}\right) \quad ,$$

$det\left(\boldsymbol{A}\right) \;\neq\; 0, \;\; \Rightarrow \boldsymbol{A}$ - невырожденная.

Векторные нормы

Норма это метрика или расстояние в линейном пространстве. Линейное пространство + метрика в нем = метрическое пространство. Норма вектора в общем случае это функция $f \;:\; \boldsymbol{R}^{\boldsymbol{n}} \rightarrow \boldsymbol{R}$, обладающая следующими свойствами:

$$f\left(\boldsymbol{x}\right)\geq 0, \quad \boldsymbol{x}\in\boldsymbol{R}^{\boldsymbol{n}}, \quad f\left(\boldsymbol{x}\right)=0 \;\Leftrightarrow \boldsymbol{x}=\boldsymbol{0} \quad ;$$

$$f\left(\boldsymbol{x}+\boldsymbol{y}\right) \;\leq\; f\left(\boldsymbol{x}\right)+f\left(\boldsymbol{y}\right), \quad \boldsymbol{x},\boldsymbol{y}\in\boldsymbol{R}^{\boldsymbol{n}} \quad ;$$

$$f\left(\alpha\boldsymbol{x}\right) \;=\; |\alpha| f\left(\boldsymbol{x}\right), \quad \alpha\in\boldsymbol{R},\boldsymbol{x}\in\boldsymbol{R}^{\boldsymbol{n}} \quad .$$

$f\left(\boldsymbol{x}\right)=\|\boldsymbol{x}\|$ - обозначение нормы вектора.

Наиболее используемый класс векторных норм — *p*-нормы:

$$\|\boldsymbol{x}\|_{p} \;=\; \left(\,|x_1|^{p}|x_2|^{p}+\ldots+|x_n|^{p}\,\right)^{1/p} \quad .$$

Наиболее используемые нормы это $1{,}2\,,\infty$ - нормы:

$$\|\boldsymbol{x}\|_{1} \;=\; \sum_{i=1}^{n}|x_i| \quad ;$$

$$\|\boldsymbol{x}\|_{2} \;=\; \sqrt{\sum_{i=1}^{n}|x_i|^{2}} \;=\; \left(\boldsymbol{x}^{\boldsymbol{T}}\cdot\boldsymbol{x}\right)^{1/2} \quad ;$$

$$\|\boldsymbol{x}\|_{\infty} \;=\; \max_{1\leq i\leq n}|x_i| \quad .$$

В системе Octave эти нормы можно вычислить подав команды *norm(x,1)*, *norm(x,2)*, *norm(x,’inf’)*.
Абсолютную и относительную погрешности для векторов теперь можно выразить в терминах векторных норм:

$$\varepsilon_{\mathrm{abs}} \;=\; \|\hat{\boldsymbol{x}}-\boldsymbol{x}\| \quad ,$$

$$\varepsilon_{rel} \;=\; \frac{\|\hat{\boldsymbol{x}}-\boldsymbol{x}\|}{\|\boldsymbol{x}\|}, \quad \|\boldsymbol{x}\|\neq 0 \quad .$$

Матричные нормы

Функция $f:\boldsymbol{R}^{\boldsymbol{m}\times\boldsymbol{n}}\rightarrow\boldsymbol{R}$ является матричной нормой, если

$$f\left(\boldsymbol{A}\right)\geq 0, \quad \boldsymbol{A}\in\boldsymbol{R}^{\boldsymbol{m}\times\boldsymbol{n}},\; f\left(\boldsymbol{A}\right)=0\Leftrightarrow \boldsymbol{A}=\boldsymbol{0} \quad ;$$

$$f(\boldsymbol{A}+\boldsymbol{B}) \le f(\boldsymbol{A})+f(\boldsymbol{B}), \qquad \boldsymbol{A},\boldsymbol{B} \in \boldsymbol{R}^{n\times m} \quad ;$$

$$f(\alpha \boldsymbol{A}) = |\alpha| f(\boldsymbol{A}), \qquad \alpha \in \boldsymbol{R}, \ \ \boldsymbol{A} \in \boldsymbol{R}^{m\times n} \quad .$$

Норма Фробениуса:

$$\|\boldsymbol{A}\|_F = \sqrt{\sum_{i-1}^{m} \sum_{j=1}^{n} |a_{ij}|^2} \quad .$$

Матричные p-нормы:

$$\|\boldsymbol{A}\|_p = \sup_{\boldsymbol{x}\ne\boldsymbol{0}} \frac{\|\boldsymbol{A}\,\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p} \quad .$$

Очевидно, что $\|\boldsymbol{A}\|_p$ это *p*-норма наибольшего вектора, полученного действием матрицы $\boldsymbol{A}$ на векторы единичной (в *p*-норме) длины, т. е.

$$\|\boldsymbol{A}\|_p = \sup_{\boldsymbol{x}\ne\boldsymbol{0}} \left\| \boldsymbol{A} \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_p} \right\| = \max_{\|\boldsymbol{x}\|_p=1} \|\boldsymbol{A}\,\boldsymbol{x}\|_p \quad .$$

Обусловленность СЛАУ

Пусть в уравнении (системе уравнений)

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f}$$

и матрица $\boldsymbol{A}$, и правая часть $\boldsymbol{f}$ заданы с погрешностью. Тогда и решение будет неточным, и реальная система будет такой:

$$(\boldsymbol{A} + \Delta \boldsymbol{A})(\boldsymbol{u} + \Delta \boldsymbol{u}) = (\boldsymbol{f} + \Delta \boldsymbol{f}) \quad .$$

Числом обусловленности принято считать величину

$$\mu(\boldsymbol{A}) = \|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\| \quad ,$$

заметим, что эта величина зависит также от выбора нормы.

Имеет место оценка

$$\frac{\|\Delta \boldsymbol{u}\|}{\|\boldsymbol{u}\|} \le \frac{\mu(\boldsymbol{A})}{1-\mu(\boldsymbol{A})\dfrac{\|\Delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \cdot \left(\frac{\|\Delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|} + \frac{\|\Delta \boldsymbol{f}\|}{\|\boldsymbol{f}\|} \right) \quad .$$

Если матрица $\boldsymbol{A}$ - точная ($\Delta \boldsymbol{A} = \boldsymbol{0}$), а неточность определяется только правой частью, то

$$\frac{\|\Delta \boldsymbol{u}\|}{\|\boldsymbol{u}\|} \le \mu(\boldsymbol{A}) \cdot \frac{\|\Delta \boldsymbol{f}\|}{\|\boldsymbol{f}\|} \quad .$$

Для случая $\Delta \boldsymbol{A} = \boldsymbol{0}$ используется и другой параметр обусловленности,

$$\nu(\boldsymbol{A}, \boldsymbol{f}) = \frac{\|\boldsymbol{f}\|}{\|\boldsymbol{u}\|} \cdot \|\boldsymbol{A}^{-1}\| \quad .$$

Есть и другие способы… Хорошо обусловленная система характеризуется числом обусловленности

$$1 \le \mu(\boldsymbol{A}) < 100 \quad .$$

Несложно показать, что это число не меньше единицы:

$$\mu(\boldsymbol{A}) = \|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\| \ge \|\boldsymbol{A}\,\boldsymbol{A}^{-1}\| = \|\boldsymbol{I}_{\boldsymbol{n}}\| = 1 \quad .$$

Пример:

$$\begin{cases} x_1 + 0.99\,x_2 = b1 \\ 0.99\,x_1 + x2 = b2 \end{cases} \quad ,$$

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0.99 \\ 0.99 & 1 \end{pmatrix} \quad ,$$

$$\|\boldsymbol{A}\|_1 = \|\boldsymbol{A}\|_2 = \|\boldsymbol{A}\|_F = 1.99 \quad ,$$

$$\mu = \|\boldsymbol{A}\|_2 \cdot \|\boldsymbol{A}^{-1}\|_2 = 199 \quad .$$