Матричные декомпозиции и решение систем линейных алгебраических уравнений

Мальковский Н. В.

Санкт-Петербургский Академический Университет

Матричная декомпозиция

Определение

Декомпозицией или *разложением* матрицы A называется представление A в виде некоторого произведения

$$A = A_1 A_2 \dots A_n$$

Замечание. Обычно $n = 2, 3$, в интересующих нах случаях A и A_i – квадратные матрицы.

Зачем нужно: Пусть $A = A_1 \dots A_n$, дана система

$$Ax = b$$

Если $A_1 y = b$, систему $Ax = b$ можно редуцировать до системы $A_2 \dots A_n x = y$.
Таким образом, интересны разложения с небольшим n и простой структурой A_i.

Ортогональные матрицы

Определение

Матрица $P \in \mathbb{R}^{n \times n}$ называется ортогональной, если $P^T P = PP^T = I$.

Замечание 1. Непосредственное следствие, если P ортогональна, то $P^{-1} = P^T$. *Замечание* 2. Пусть $p_1, \ldots, p_n \in K^n$ – столбцы матрицы P, т. е.

$P = [p_1 \ldots p_n]$, тогда

$$P^T P = \begin{bmatrix} p_1^T \\ \ldots \\ p_n^T \end{bmatrix} [p_1 \ldots p_n] = \begin{pmatrix} p_1^T p_1 & p_1^T p_2 & \ldots & p_1^T p_n \\ p_2^T p_1 & p_2^T p_2 & \ldots & p_2^T p_n \\ \vdots & \vdots & \ddots & \vdots \\ p_n^T p_1 & p_n^T p_2 & \ldots & p_n^T p_n \end{pmatrix} = I$$

Таким образом

$$p_i^T p_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

иначе говоря, столбцы P – ортонормированы. Это свойство выполняется и для строк.

Треугольные матрицы

Определение

Матрица $A = (a_{ij}) \in K^{n \times n}$ *называется* *верхней треугольной* *матрицей, если* $a_{ij} = 0$ *при* $i > j$. *Аналогично* A *называется* *нижней треугольной*, *если* $a_{ij} = 0$ *при* $i < j$.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} \qquad \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Решение СЛАУ для ортогональных матриц

Пусть P – ортогональная матрица, рассмотрим систему

$$Px = b$$

Из свойств P получаем

$$x = P^T P x = P^T b$$

Фактически, свойство ортогональности можно переписать в виде $P^{-1} = P^T$.

Треугольные матрицы

Свойства:

1. Произведение верхних треугольных матриц есть верхняя треугольная матрица, при $i > j$

$$\sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{j} \underbrace{a_{ik}}_{=0} b_{kj} + \sum_{k=j+1}^{n} a_{ik} \underbrace{b_{kj}}_{=0} = 0$$

2. Пусть $A = (a_{ij})$ – верхняя треугольная матрица с ненулевыми диагональными элементами, тогда A^{-1} существует и также является верхней треугольной матрицей. Пусть $i = n$, тогда для $j < n$ имеем

$$\sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n-1} \underbrace{a_{ik}}_{=0} b_{kj} + a_{nn} b_{nj} = 0 \;\Rightarrow\; b_{nj} = 0$$

Повторяя для $i = n - 1 \dots j + 1$ получаем

$$\sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{i-1} \underbrace{a_{ik}}_{=0} b_{kj} + a_{ii} b_{ij} + \sum_{k=i+1}^{n} a_{ik} \underbrace{b_{kj}}_{=0,\ \text{индукция}} = 0 \;\Rightarrow\; b_{ij} = 0$$

Треугольные матрицы

3. Если A – верхняя треугольная, то $A-\lambda I$ тоже верхняя треугольная матрица, а её определитель легко вычисляется

$$\det(A-\lambda I)=\prod_{i=1}^{n}(a_{ii}-\lambda)$$

Таким образом, диагональные элементы A являются её собственными числами с учетом кратности.

4. Атомарная треугольная матрица – нижняя треугольная матрица следующего вида

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & a_{(i+1)i} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{ni} & 0 & \dots & 1 \end{pmatrix}$$

Домножение слева на такую матрицу добавляет к строке i линейную комбинацию строк $i+1,\dots,n$.

5. Обратная матрица к атомарной вычисляется довольно просто

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & a_{(i+1)i} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{ni} & 0 & \dots & 1 \end{pmatrix} \quad A^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & -a_{(i+1)i} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -a_{ni} & 0 & \dots & 1 \end{pmatrix}$$

Решение СЛАУ для треугольных матриц

Пусть L – нижняя треугольная матрица, тогда система $Lx = b$ имеет вид

$$\begin{array}{llllll}
\ell_{11}x_1 & & & & = & b_1 \\
\ell_{21}x_1 & + \ell_{22}x_2 & & & = & b_2 \\
\vdots & \vdots & \ddots & & = & \vdots \\
\ell_{n1}x_1 & + \ell_{n2}x_n & \dots & + \ell_{nn}x_n & = & b_n
\end{array}$$

Из-за особого вида системы решение легче всего получить последовательным решением уравнений в заданном порядке, что дает рекуррентные формулы

$$x_i = \frac{b_i - \sum_{k=1}^{i-1} \ell_{ik}x_k}{\ell_{ii}}$$

Замечание 1. Если все диагональные элементы ненулевые, то формулы корректны. Если же $\ell_{ii} = 0$, но при этом $b_i - \sum_{k=1}^{i-1} \ell_{ik}x_k \neq 0$, то решения не существует, если же $b_i - \sum_{k=1}^{i-1} \ell_{ik}x_k$, то x_i можно выбрать любое. *Замечание* 2. Нахождение решения по этим формулам принято называть прямой подстановкой.

Решение СЛАУ для треугольных матриц

Пусть L – нижняя треугольная матрица, тогда система $Lx = b$ имеет вид

$$\begin{array}{cccccc} u_{11}x_1 & +\, u_{12}x_n & \dots & +\, u_{1n}x_n & = & b_1 \\ & \ddots & \vdots & \vdots & = & \vdots \\ & & u_{(n-1)(n-1)}x_{n-1} & +u_{(n-1)n}x_n & = & b_{n-1} \\ & & & u_{nn}x_n & = & b_n \end{array}$$

Эту систему решать проще в обратном порядке, по аналогии с верхней треугольной матрицей получаем следующие формулы

$$x_i = \frac{b_i - \sum_{k=i+1}^{n} u_{ik}x_k}{u_{ii}}$$

Замечание. Такой процесс принято называть обратной подстановкой.

LU-декомпозиция

Определение

LU-декомпозицией матрицы $A \in K^{n \times n}$ называется представление $A = LU$, где $L, U \in K^{n \times n}$, L – нижняя диагональная матрица, а U – верхняя диагональная матрица.

Замечание. Уравнение $Ax = b$ переписывается в виде $L^{-1}Ax = Ux = L^{-1}b$. Метод Гаусса заключается в последовательном домножении левой и правой части исходного уравнения на атомарные матрицы, что в итоге дает приведенное уравнение $Ux = L^{-1}b$.

Метод Гаусса

Метод Гаусса заключается в последовательном вычитанием одного уравнения из других, таким образом множество решений системы не меняется. Метод делится на n шагов, на шаге k система имеет вид

$$\begin{pmatrix} u_{11}^k & u_{12}^k & \dots & u_{1k}^k & \dots & u_{1n}^k \\ 0 & u_{22}^k & \dots & u_{2k}^k & \dots & u_{2n}^k \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{kk}^k & \dots & u_{kn}^k \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nk}^k & \dots & u_{nn}^k \end{pmatrix} x = \begin{pmatrix} b_1^k \\ b_2^k \\ \vdots \\ b_k^k \\ \vdots \\ b_n^k \end{pmatrix}$$

Здесь под u_{ij}^k подразумевается значение соответствующего коэффициента на шаге k. $u_{ij}^k = 0$ при выполнении двух условий:

- $i > j$
- $j \leq k$

Метод Гаусса

Итерация k: вычесть промасштабированную строку k из строк $k+1,\ldots,n$ так, чтобы обнулить столбец k.

Непосредственным вычислением получаем для:

$$u_{ij}^{k} = \begin{cases} u_{ij}^{k-1} - \frac{u_{kk}^{k-1}}{u_{ik}^{k-1}} u_{kj}, & i > k \\ u_{ij}^{k-1} & i \leq k \end{cases} \quad b_{i}^{k} = \begin{cases} b_{i}^{k-1} - \frac{u_{kk}^{k-1}}{u_{ik}^{k-1}} b_{k}^{k-1}, & i > k \\ b_{i}^{k-1}, & i \leq k \end{cases}$$

Итерация k требует

- $n-k$ делений.
- $(n-k)^2-1$ сложений и умножений.

Замечание. Если A обратима, то все формулы корректны и u_{kk}^{k-1} отлично от нуля.

Пример

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} U = \begin{pmatrix} 2 & 3 & -2 & -1 & 0 \\ -2 & -2 & 4 & 4 & 1 \\ -4 & -7 & 0 & -1 & -1 \\ 2 & 5 & 0 & 6 & 1 \\ -2 & -5 & -4 & -4 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 & -2 & -1 & 0 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -1 & -4 & -3 & -1 \\ 0 & 2 & 2 & 7 & 1 \\ 0 & -2 & -6 & -5 & 0 \end{pmatrix}$$

$$A = \begin{pmatrix} 2 & 3 & -2 & -1 & 0 \\ -2 & -2 & 4 & 4 & 1 \\ -4 & -7 & 0 & -1 & -1 \\ 2 & 5 & 0 & 6 & 1 \\ -2 & -5 & -4 & -4 & 0 \end{pmatrix} L_i = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ \ldots \end{pmatrix} \begin{pmatrix} 2 & 3 & -2 & -1 & 0 \\ 0 & 1 & 2 & 3 & 1 \\ \ldots \end{pmatrix}$$

Рекурсивная форма метода Гаусса

Пусть $A \in K^{n \times n}$, $\ell \in K^{n-1}$, $u \in K^{n-1}$, $\tilde{A} \in K^{n-1 \times n-1}$ и при этом

$$A = \left[\begin{array}{cc} a_{11} & u^T \\ \ell & \tilde{A} \end{array} \right]$$

Обозначим за $GU(A)$, $GL(A)$ – верхнюю и нижнюю треугольную матрицу, получающуюся в результате применения метода Гаусса к матрице A, тогда имеют место соотношения

$$GU(A) = \left[\begin{array}{cc} a_{11} & u^T \\ 0 & GU(\tilde{A} - \frac{1}{a_{11}} \ell u^T) \end{array} \right]$$

и

$$GL(A) = \left[\begin{array}{cc} 1 & 0_{n-1}^T \\ -\frac{1}{a_{11}} \ell & I_{n-1} \end{array} \right] \left[\begin{array}{cc} 1 & 0_{n-1}^T \\ 0_{n-1} & GL(\tilde{A} - \frac{1}{a_{11}} \ell u^T) \end{array} \right]$$

Разложение Холеского

Определение

Разложением Холеского называется представление квадратной матрицы в виде $A = LL^T$, где L – нижняя треугольная матрица.

Теорема

Для квадратной матрицы A разложение Холеского существует тогда и только тогда, когда она симметрична и положительно определена.

Док-во. Необходимость: легко увидеть, что матрица LL^T всегда симметрична и положительно определена

$$x^T LL^T x = (L^T x)^T L^T x = ||L^T x||^2 \geq 0$$

Разложение Холеского

Достаточность: пусть A – симметричная положительно определенная матрица и пусть, $A = LU$, где L – обратимая нижняя треугольная матрица, а U – верхняя треугольная матрица (существует в следствии метода Гаусса), тогда

$$A = A^T = LU = U^T L^T$$

$$U(L^T)^{-1} = L^{-1} U^T.$$

В левой части последнего равенства находится произведение верхних треугольных матриц, в правой – произведение нижних треугольных матриц. Равенство возможно только в том случае, когда в обоих частях диагональные матрицы.

Обозначим $D = \mathrm{diag}\{d_1, \ldots, d_n\} = U(L^T)^{-1}$, тогда

$$A = LDL^T$$

В силу обратимости L и положительной определенности A получаем положительную определенность D, т. е. $d_i \geq 0$, отсюда $A = L\sqrt{D}(L\sqrt{D})^T$, где $\sqrt{D} = \mathrm{diag}\{\sqrt{d_1}, \ldots, \sqrt{d_n}\}$.

Разложение Холеского

Замечание 1. Если матрица A строго положительно определена, то разложение Холеского единственно.
Расписывая равенство
$$LL^T = A$$
покоординатно получаем получаем для $i \geq j$
$$a_{ij} = \sum_{k=1}^{n} \ell_{ik}\ell_{jk} = \sum_{k=1}^{j} \ell_{ik}\ell_{jk}$$
Из этих равенств можно получаем следующие рекуррентные формулы
$$\begin{aligned} \ell_{ii} &= \pm\sqrt{a_{ii} - \sum_{j=1}^{i-1} \ell_{ij}^2} \\ \ell_{ij} &= \frac{1}{\ell_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} \ell_{ik}\ell_{jk} \right), \quad j < i \end{aligned} \tag{1}$$

Разложение Холеского

Если A положительно определена, то выражение корнем всегда положительно, так как система должна иметь решение. В случае, если $\ell_{jj} = 0$, то ℓ_{ij} может быть выбрано произвольно.

Пусть λ – минимальное собственное число A соответственно, а γ – минимальное собственное число L соответственно. Для этих величин очевидным образом выполняются следующие соотношения:

$$\lambda = \min_{||x||=1} x^T A x = \min_{||x||=1} ||L^T x||^2 \geq \gamma^2$$

Таким образом L содержит собственное число 0 тогда и только тогда когда A содержит A. Следовательно для строго положительно определенной A формулы (1) задают единственную матрицу L с точностью до знака перед корнем. Итого имеем следующую теорему единственности

Теорема

Для строго положительно определенной матрицы A существует единственная строго положительно определенная нижняя треугольная матрица L такая, что $A = LL^T$.

Рекурсивная форма разложения Холеского

Пусть $A \in K^{n \times n}$ – симметричная строго положительно определенная матрица, $\ell \in K^{n-1}$, $\tilde{A} \in K^{n-1 \times n-1}$ и при этом

$$A = \begin{bmatrix} a_{11} & \ell^T \\ \ell & \tilde{A} \end{bmatrix}$$

Отметим, что имеет место равенство

$$\begin{bmatrix} a_{11} & \ell^T \\ \ell & \tilde{A} \end{bmatrix} = \begin{bmatrix} \sqrt{a_{11}} & 0_{n-1}^T \\ \frac{1}{\sqrt{a_{11}}}\ell & I_{n-1} \end{bmatrix} \begin{bmatrix} 1 & 0_{n-1}^T \\ 0_{n-1} & \tilde{A} - \frac{1}{a_{11}}\ell\ell^T \end{bmatrix} \begin{bmatrix} \sqrt{a_{11}} & \frac{1}{\sqrt{a_{11}}}\ell^T \\ 0_{n-1} & I_{n-1} \end{bmatrix}$$

Таким образом, обозначив за $HL(A)$ – единственную матрицу L, что $A = LL^T$ и используя факт, что при $B = RR^T$ выполняется

$$\begin{bmatrix} I & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & R^T \end{bmatrix}$$

получаем

$$HL(A) = \begin{bmatrix} \sqrt{a_{11}} & 0_{n-1}^T \\ \frac{1}{\sqrt{a_{11}}}\ell & HL(\tilde{A} - \frac{1}{a_{11}}\ell u^T) \end{bmatrix}$$

Итеративные схемы

Рассмотрим снова СЛАУ
$$Ax = b.$$
Если D – обратимая матрица, то система может быть эквивалентна переписана в виде
$$x = x + D(Ax - b) = (I - DA)x + Db$$
Общая <u>итеративная схема решения СЛАУ</u> заключается в построении последовательности
$$x_{k+1} = (I - DA)x_k + Db \tag{2}$$

Теорема

Если матрица $I - DA$ положительно определена, $\sigma(I - DA) < 1$ и D обратима, то (2) сходится к сходится решению системы $Ax = b$.

Док-во. Сходимость и скорость сходимости была показана раньше (сходимость линейных итеративных процессов). Пусть $x_k \to x^*$, тогда $x^* = x^* - D(Ax^* - b)$, т. е. $D(Ax^* - b) = 0$, из обратимости D следует $Ax^* = b$. ∎

Два способа построения итеративных схем

Основная проблема обычно заключается в том, чтобы подобрать D так, что $\sigma(I - DA) < 1$. Два наиболее распространенных случая:

1. $0 \preceq A \preceq \gamma I$, тогда годится выбор $D = \gamma^{-1} I$. Для симметричной матрицы A этот метод идентичен применению градиентного спуску к функции

$$f(x) = \frac{1}{2} x^T A x + b^T x + c$$

2. A имеет диагональное преобладание, т. е.

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

   В этом случае годится выбор $D = \operatorname{diag}\{a_{11}^{-1}, \ldots, a_{nn}^{-1}\}$. Этот метод принято называть *методом Якоби*.

Теорема (Гершгорина о кругах)

Пусть $A = (a_{ij}) \in \mathbb{C}^{n\times n}$. Для любого собственного числа λ матрицы A найдется $1 \le i \le n$ такое, что

$$|\lambda - a_{ii}| \le \sum_{j\neq i} |a_{ij}|$$

Док-во. Пусть $Ax = \lambda x$ и $i = \operatorname{argmax}_j |x_j|$. Таким образом

$$\sum_j a_{ij} x_j = \lambda x_i$$

Поделив обе части на x_i получаем

$$\sum_{j\neq i} a_{ij}\frac{x_j}{x_i} + a_{ii} = \lambda$$

Таким образом

$$|\lambda - a_{ii}| = \left|\sum_{j\neq i} a_{ij}\frac{x_j}{x_i}\right| \le \sum_{j\neq i} |a_{ij}| \quad \blacksquare$$

Теорема Гершгорина

Замечание 1. Вещественная часть любого собственного числа вещественной матрицы с диагональным преобладанием и строго положительными диагональными элементами неотрицательна.

Замечание 2. В методе Якоби выполняется $0 \preceq (I - DA) \preceq I$.

Метод сопряженных градиентов

Определение

Пусть A – симметричная матрица. Вектора u, v называются <u>A-ортогональными</u> или <u>сопряженными</u>, если

$$u^T A v = 0$$

Рассмотрим задачу минимизации

$$f(x) = \frac{1}{2} x^T A x - b^T x + c,$$

где A – симметричная обратимая матрица, таким образом $\nabla f(x) = Ax - b$, а значит нахождение точки минимума f равносильно решению системы $Ax = b$. Обозначим за x^* единственную точку минимума f.

Предположим, что нам известны n попарно сопряженных направлений $d_1, \ldots, d_n$. Выберем произвольную точку x_0 и сделаем по очереди n шагов градиентного спуска по каждому из направлений, выбирая размер шага как минимум по направлению.

Метод сопряженных градиентов

Таким образом получаем

$$x_{k+1} = x_k - \alpha_k d_k$$

Получаем α_k из уравнения $\frac{d}{d\alpha} f(x_k - \alpha d_k) = 0$:

$$0 = \frac{d}{d\alpha} f(x_k - \alpha d_k) = d_k^T (A(x_k - \alpha d_k) - b) = -\alpha d_k^T A d_k^T + d_k^T (Ax_k - b)$$

Обозначим $r_k = Ax_k - b$:

$$\alpha_k = \frac{d_k^T r_k}{d_k^T A d_k}$$

Метод сопряженных градиентов

Теперь предположим, что $(d_0, \ldots, d_{n-1})$ – базис в $\mathbb{R}^n$, тогда

$$x_0 - x^* = \sum_{i=0}^{n-1} \delta_i d_i$$

Умножая это равенство на $d_k^T A$ получаем

$$d_k^T A(x_0 - x^*) = \sum_{i=0}^{n-1} \delta_i d_k^T A d_i = \delta_k d_k^T A d_k$$

и получаем следующие равенства для δ

$$\begin{aligned}\delta_k &= \frac{d_k^T A(x_0 - x^*)}{d_k^T A d_k} \\ &= \frac{d_k^T A(x_0 - x^* - \sum_{i=0}^{k-1} \alpha_i d_i)}{d_k^T A d_k} = \frac{d_k^T A(x_k - x^*)}{d_k^T A d_k} = \alpha_k\end{aligned}$$

Метод сопряженных градиентов

Таким образом получаем

$$x_k - x^* = \sum_{i=k}^{n-1} \alpha_i d_i,$$

что гарантирует сходимость этой процедуры за n шагов.

Остается ответить на вопрос: как найти n сопряженных векторов?

Модификация процедуры Грама-Шмидта

Пусть $v_1, \ldots, v_n \in \mathbb{R}^n$, рассмотрим $u_1, \ldots, u_n$:

$$u_k = v_k - \sum_{i=1}^{k-1} \frac{u_i^T A v_k}{u_i^T A u_i} u_i$$

Аналогично обычной процедуре Грама-Шмидта докажем по индукции, что

$u_j^T A u_k = 0$ при $j \neq k$. Пусть это утверждение верно до $k-1$, тогда для $j < k$

$$u_j^T A u_k = u_j^T A v_k - \sum_{i=1}^{k-1} \frac{u_i^T A v_k}{u_i^T A u_i} u_j^T A u_i = u_j^T A v_k - u_j^T A v_k = 0.$$

Также по индукции доказывается равенство линейных оболочек $\mathrm{Span}\{v_1, \ldots, v_k\} = \mathrm{Span}\{u_1, \ldots, u_k\}$.

Общая схема метода сопряженных направлений

Итого, на данный момент имеем следующую общую схему:

1. Выбрать n линейно независимых векторов $v_1, \dots, v_n$.
2. Построить n сопряженных относительно матрицы A направлений $d_1, \dots, d_n$ по формулам

$$d_k = v_k - \sum_{i=1}^{k-1} \frac{d_i^T A v_k}{d_i^T A d_i} d_i$$

3. Выбрать произвольную точку x_0 и построить последовательность

$$x_{k+1} = x_k - \frac{d_k^T (A x_k - b)}{d_k^T A d_k} d_k$$

Из показанного ранее вытекает, что $Ax_n = b$. Если m – количество ненулевых элементов A, то умножение A на вектор требует $\mathcal{O}(m)$ сложений и умножений. Таким образом, шаг 2 требует суммарно $\mathcal{O}(n^2 m)$ действий, шаг 3 – $\mathcal{O}(nm)$.

Сопряженные градиенты

Основная идея метода сопряженных градиентов – использование $v_k = Ax_k - b$.

$$v_i^T(Ax_{k+1} - b) = v_i^T(Ax_k - b) - \alpha_k v_i^T A d_k$$

$$v_i^T A d_k = \frac{1}{\alpha_k}\left(v_i^T v_k - v_i^T v_{k+1}^T\right)$$

Пусть $\mathcal{D}_i = \text{Span}\{d_0, Ad_0, \ldots, A^{i-1}d_0\}$, тогда $d_{i-1} \in \mathcal{D}_i$. В силу $d_k^T A(x_j - x^*) = \sum_{i=k}^{n-1} \alpha_i d_k^T A d_i$ получаем ортогональность $d_k^T v_j = 0$ при $k < j$. Итого шаг 2 имеет вид

$$d_{k+1} = v_{k+1} + \frac{v_{k+1}^T v_{k+1}}{v_k^T v_k} d_k$$