О МЕТОДЕ СОПРЯЖЁННЫХ ГРАДИЕНТОВ*

В. Н. Малозёмов
malv@math.spbu.ru

28 апреля 2012 г.

В докладе речь пойдёт об одном эффектном методе минимизации квадратичной функции, предложенном в начале 1950-х годов Хестинсом и Штифелем [1]. Будут использоваться некоторые сведения из книги [2, с. 433–454].

1°. Сопряжённые направления. Пусть D — симметричная положительно определённая матрица порядка n. Два ненулевых вектора s_1, s_2 из $\mathbb{R}^n$ называются D-ортогональными или *сопряжёнными*, если

$$\langle s_2, Ds_1 \rangle = \langle Ds_2, s_1 \rangle = 0.$$

Рис. 1 поясняет это определение.

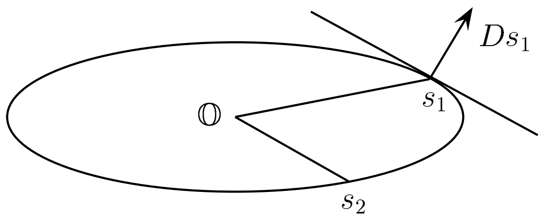

Рис. 1

Опишем динамический метод построения последовательности D-ортогональных векторов. Возьмём произвольный ненулевой вектор $y_1 \in \mathbb{R}^n$ и положим $s_1 = y_1$.

Пусть $y_2 \in \mathbb{R}^n$ — произвольный ненулевой вектор, линейно независимый с y_1. Сопряжённый к s_1 вектор s_2 будем искать в виде

$$s_2 = y_2 + \gamma_{21} s_1.$$

*Семинар по дискретному гармоническому анализу и геометрическому моделированию «DHA & CAGD»: `http://www.dha.spb.ru/`

Отметим, что $s_2 = y_2 + \gamma_{21} y_1$. В силу линейной независимости y_1 и y_2 вектор s_2 отличен от нулевого при любом γ_{21}. Найдём γ_{21} из условия D-ортогональности $\langle s_2, Ds_1 \rangle = 0$. Получим

$$\gamma_{21} = -\frac{\langle y_2, Ds_1 \rangle}{\langle s_1, Ds_1 \rangle}.$$

Продолжим процесс. Возьмём произвольный ненулевой вектор $y_3 \in \mathbb{R}^n$, линейно независимый с y_1 и y_2. Очередное сопряжённое направление s_3 будем искать в виде

$$s_3 = y_3 + \gamma_{31} s_1 + \gamma_{32} s_2.$$

Учитывая, что $s_3 = y_3 + \alpha_{31} y_1 + \alpha_{32} y_2$, в силу линейной независимости векторов y_1, y_2, y_3 заключаем, что $s_3 \neq \mathbb{O}$ при любых γ_{31}, γ_{32}. Коэффициенты γ_{31} и γ_{32} найдём из условий D-ортогональности

$$\langle s_3, Ds_1 \rangle = 0, \quad \langle s_3, Ds_2 \rangle = 0.$$

Получим

$$\gamma_{31} = -\frac{\langle y_3, Ds_1 \rangle}{\langle s_1, Ds_1 \rangle}, \quad \gamma_{32} = -\frac{\langle y_3, Ds_2 \rangle}{\langle s_2, Ds_2 \rangle}.$$

Пусть уже построены D-ортогональные векторы $s_1, s_2, \ldots, s_{k-1}$ с помощью последовательно привлекаемых линейно независимых векторов $y_1, y_2, \ldots, y_{k-1}$. Возьмём произвольный ненулевой вектор $y_k \in \mathbb{R}^n$, линейно независимый с $y_1, y_2, \ldots, y_{k-1}$. Сопряжённое направление s_k будем искать в виде

$$s_k = y_k + \sum_{j=1}^{k-1} \gamma_{kj} s_j. \tag{1}$$

В силу линейной независимости векторов y_1, y_2, $\ldots$, y_k вектор s_k отличен от нулевого при любых γ_{kj}. Найдём коэффициенты γ_{kj} из условий D-ортогональности: $\langle s_k, Ds_i \rangle = 0$ при всех $i \in 1 : k-1$. Получим

$$\gamma_{kj} = -\frac{\langle y_k, Ds_j \rangle}{\langle s_j, Ds_j \rangle}, \quad i \in 1 : k-1. \tag{2}$$

При $k = n$ будет построена полная в $\mathbb{R}^n$ система D-ортогональных векторов $s_1, s_2, \ldots, s_n$.

Эти векторы линейно независимы. Действительно, если

$$\sum_{j=1}^{n} c_j s_j = \mathbb{O},$$

то после умножения обеих частей этого равенства скалярно на Ds_i получим $c_i = 0$ при всех $i \in 1 : n$.

Обозначим через S матрицу со столбцами $s_1, s_2, \ldots, s_n$. В силу D-ортогональности

$$S^T(DS) = \Lambda, \tag{3}$$

где Λ — диагональная матрица с диагональными элементами $\Lambda[i,i] = \langle Ds_i, s_i \rangle$, $i \in 1:n$. Из (3) следует, что $D = (S^T)^{-1}\Lambda S^{-1}$ и $D^{-1} = S(\Lambda^{-1}S^T)$. Последнее равенство можно записать в виде

$$D^{-1} = \sum_{i=1}^{n} \frac{s_i s_i^T}{\langle Ds_i, s_i \rangle}. \tag{4}$$

2°. Обратимся к задаче минимизации квадратичной функции:

$$Q(x) := \tfrac{1}{2}\langle Dx, x \rangle + \langle c, x \rangle \to \min_{x \in \mathbb{R}^n}. \tag{5}$$

Будем считать, что матрица D симметрична и положительно определена.

Напомним, что для градиента квадратичной функции справедлива формула

$$Q'(x) = Dx + c.$$

Условие $Q'(x) = \mathbb{O}$ служит критерием оптимальности для задачи (5). Значит, решение экстремальной задачи (5) равносильно решению системы линейных уравнений

$$Dx = -c. \tag{6}$$

Это принципиальный факт.

Система (6) имеет единственное решение

$$x_* = -D^{-1}c. \tag{7}$$

Этот же вектор x_* является единственным решением задачи (5).

Предположим, что построена какая-нибудь система D-ортогональных векторов $s_1, s_2, \ldots, s_n$. Возьмём произвольный вектор $x_0 \in \mathbb{R}^n$ и вычислим $g_0 := Q'(x_0) = Dx_0 + c$. На основании (7) и (4) получим

$$\begin{aligned} x_* = x_0 - D^{-1}(Dx_0 + c) = x_0 - D^{-1}g_0 = \\ = x_0 - \sum_{i=1}^{n} \frac{\langle s_i, g_0 \rangle}{\langle Ds_i, s_i \rangle} s_i = x_0 + \sum_{i=1}^{n} t_i s_i, \end{aligned}$$

где

$$t_i = -\frac{\langle g_0, s_i \rangle}{\langle Ds_i, s_i \rangle}.$$

Обозначим

$$x_k = x_0 + \sum_{i=1}^{k} t_i s_i.$$

Очевидно, что

$$x_k = x_{k-1} + t_k s_k, \quad k = 1, \dots, n. \tag{8}$$

При этом $x_n = x_*$.

Приходим к следующему выводу: *при любом выборе D-ортогональной системы векторов $s_1, s_2, \dots, s_n$ (точнее, при любом выборе линейно независимой системы векторов $y_1, y_2, \dots, y_n$) и произвольном начальном приближении x_0 вычисления по рекуррентной формуле* (8) *приводят к единственному решению задачи* (5).

Положим $g_k := Q'(x_k) = Dx_k + c$. Согласно (8)

$$g_k - g_{k-1} = t_k D s_k, \quad k \in 1 : n. \tag{9}$$

Отметим одну особенность метода сопряжённых направлений.

ПРЕДЛОЖЕНИЕ 1. *Справедливо равенство*

$$\langle g_k, s_i \rangle = 0, \quad i \in 1 : k. \tag{10}$$

Д о к а з а т е л ь с т в о. Воспользуемся формулой (9). Получим

$$\langle g_k, s_i \rangle = \big\langle (g_k - g_{k-1}) + (g_{k-1} - g_{k-2}) + \dots + (g_1 - g_0) + g_0, s_i \big\rangle =$$
$$= \langle \sum_{j=1}^{k} t_j D s_j + g_0, s_i \rangle = t_i \langle D s_i, s_i \rangle + \langle g_0, s_i \rangle = 0. \qquad \square$$

Умножим обе части равенства (9) скалярно на s_k. Приняв во внимание, что $\langle g_k, s_k \rangle = 0$, придём к основному представлению для коэффициентов t_k:

$$t_k = -\frac{\langle g_{k-1}, s_k \rangle}{\langle D s_k, s_k \rangle}, \quad k \in 1 : n. \tag{11}$$

Коэффициент t_k обладает экстремальным свойством, а именно, минимум функции $Q(x_{k-1} + t s_k)$ достигается при $t = t_k$. Это следует из (11) и разложения квадратичной функции

$$Q(x_{k-1} + t s_k) = Q(x_{k-1}) + t \langle g_{k-1}, s_k \rangle + \tfrac{1}{2} t^2 \langle D s_k, s_k \rangle. \tag{12}$$

Отметим, что разложение (12) при $t = t_k$ преобразуется к виду

$$Q(x_{k-1}) - Q(x_k) = \frac{1}{2} \frac{\langle g_{k-1}, s_k \rangle^2}{\langle D s_k, s_k \rangle}.$$

Указанное экстремальное свойство коэффициентов t_k допускает усиление.

ПРЕДЛОЖЕНИЕ 2. *Минимум квадратичной функции $Q(x)$ на аффинном множестве*

$$\mathcal{L} = \{x = x_0 + \sum_{i=1}^{k} z_i s_i\}$$

достигается только при $x = x_k$ (то есть при $z_i = t_i$, $i \in 1:k$).

Доказательство. Введём функцию

$$q_k(z) := Q(x_0 + \sum_{i=1}^{k} z_i s_i) = Q(x_0) + \sum_{i=1}^{k} \langle g_0, s_i\rangle z_i + \tfrac{1}{2} \sum_{i=1}^{k} \langle Ds_i, s_i\rangle z_i^2.$$

Видим, что $q_k(z)$ — квадратичная функция с диагональной матрицей квадратичной формы D_k, у которой $D_k[i,i] = \langle Ds_i, s_i\rangle$, $i \in 1:n$. Матрица D_k является симметричной и положительно определённой. Критерием оптимальности для экстремальной задачи

$$q_k(z) \to \min_{z \in \mathbb{R}^k}$$

служит условие $\langle Ds_i, s_i\rangle z_i = -\langle g_0, s_i\rangle$, $i \in 1:k$. Для компонент оптимального вектора z получаем формулу

$$z_i = -\frac{\langle g_0, s_i\rangle}{\langle Ds_i, s_i\rangle} = t_i, \quad i \in 1:k.$$

Это и требовалось доказать. □

3°. Метод сопряжённых градиентов. Рассмотрим конкретную последовательность привлекаемых к D-ортогонализации векторов и, тем самым, конкретный вариант метода сопряжённых направлений.

Возьмём произвольное начальное приближение $x_0 \in \mathbb{R}^n$ и вычислим градиент $g_0 = Dx_0 + c$. Если $g_0 = \mathbb{O}$, то x_0 — решение задачи (5). Процесс закончен.

Пусть $g_0 \neq \mathbb{O}$. Положим $s_1 = -g_0$ (то есть $y_1 = -g_0$). По общей схеме вычисляем

$$x_1 = x_0 + t_1 s_1,$$

где

$$t_1 = -\frac{\langle g_0, s_1\rangle}{\langle Ds_1, s_1\rangle} = \frac{\langle g_0, g_0\rangle}{\langle Ds_1, s_1\rangle}.$$

Пересчитаем градиент $g_1 = g_0 + t_1 D s_1$. Если $g_1 = \mathbb{O}$, то x_1 — решение задачи (5). Процесс закончен.

Пусть $g_1 \neq \mathbb{O}$. Отметим, что

$$\langle g_1, g_0\rangle = \langle g_0, g_0\rangle - t_1 \langle Ds_1, s_1\rangle = 0.$$

Таким образом, вектор $y_2 = -g_1$ ортогонален y_1. Его можно привлечь к процессу D-ортогонализации. Полагаем

$$s_2 = -g_1 + \gamma_{21} s_1,$$

где

$$\gamma_{21} = -\frac{\langle y_2, Ds_1\rangle}{\langle s_1, Ds_1\rangle} = \frac{\langle g_1, g_1 - g_0\rangle}{t_1\langle Ds_1, s_1\rangle} = \frac{\langle g_1, g_1\rangle}{\langle g_0, g_0\rangle} =: b_1.$$

Приходим к формуле

$$s_2 = -g_1 + b_1 s_1.$$

Предположим, что уже построены x_{k-1}, $g_{k-1} \neq \mathbb{O}$, s_k. При этом градиенты $g_0, g_1, \ldots, g_{k-1}$ попарно ортогональны, при $i \in 1 : k-1$

$$s_{i+1} = -g_i + b_i s_i, \qquad b_i = \frac{\langle g_i, g_i\rangle}{\langle g_{i-1}, g_{i-1}\rangle}, \tag{13}$$

и, по общему свойству метода сопряжённых направлений,

$$\langle g_{k-1}, s_i\rangle = 0, \quad i \in 1 : k-1. \tag{14}$$

Находим очередное приближение

$$x_k = x_{k-1} + t_k s_k,$$

где

$$t_k = -\frac{\langle g_{k-1}, s_k\rangle}{\langle Ds_k, s_k\rangle} = -\frac{\langle g_{k-1}, -g_{k-1} + b_{k-1}s_{k-1}\rangle}{\langle Ds_k, s_k\rangle} = \frac{\langle g_{k-1}, g_{k-1}\rangle}{\langle Ds_k, s_k\rangle}.$$

Пересчитываем градиент $g_k = g_{k-1} + t_k D s_k$. Если $g_k = \mathbb{O}$, то x_k — решение задачи (5). Процесс закончен.

Пусть $g_k \neq \mathbb{O}$. Покажем, что

$$\langle g_k, g_i\rangle = 0, \quad i \in 0 : k-1. \tag{15}$$

В силу D-ортогональности и формул (13), (14) при $i \in 0 : k-2$ имеем (считаем, что $s_0 = \mathbb{O}$)

$$\langle g_k, g_i\rangle = \langle g_{k-1} + t_k D s_k, -s_{i+1} + b_i s_i\rangle = 0.$$

К этому нужно добавить, что

$$\begin{aligned}\langle g_k, g_{k-1}\rangle &= \langle g_{k-1} + t_k D s_k, -s_k + b_{k-1} s_{k-1}\rangle = \\ &= -\langle g_{k-1}, s_k\rangle - t_k\langle Ds_k, s_k\rangle = 0.\end{aligned}$$

Соотношение (15) установлено.

Вектор $y_{k+1} = -g_k$ привлекаем к процессу D-ортогонализации. Запишем

$$s_{k+1} = -g_k + \sum_{i=1}^{k} \gamma_{k+1,i} s_i.$$

Здесь

$$\gamma_{k+1,i} = -\frac{\langle y_{k+1}, Ds_i \rangle}{\langle s_i, Ds_i \rangle} = \frac{\langle g_k, g_i - g_{i-1} \rangle}{t_i \langle Ds_i, s_i \rangle}.$$

Согласно (15) имеем $\gamma_{k+1,i} = 0$ при $i \in 1 : k-1$ и

$$\gamma_{k+1,k} = \frac{\langle g_k, g_k \rangle}{\langle g_{k-1}, g_{k-1} \rangle} =: b_k.$$

Таким образом,

$$s_{k+1} = -g_k + b_k s_k.$$

Описание метода завершено. Он называется *методом сопряжённых градиентов*, поскольку именно текущие градиенты привлекаются к процессу D-ортогонализации.

4°. Вычислительная схема метода сопряжённых градиентов.

Нулевой шаг. Берём произвольное начальное приближение $x_0 \in \mathbb{R}^n$ и вычисляем градиент $g_0 = Dx_0 + c$. Если $g_0 = \mathbb{O}$, то x_0 — решение задачи (5). Вычисления прекращаются. Иначе полагаем $s_1 = -g_0$.

k-й шаг. Пусть уже имеются x_{k-1}, $g_{k-1} \neq \mathbb{O}$ и s_k. Последовательно вычисляем

$$\begin{aligned} t_k &= \frac{\langle g_{k-1}, g_{k-1} \rangle}{\langle Ds_k, s_k \rangle}, \\ x_k &= x_{k-1} + t_k s_k, \\ g_k &= g_{k-1} + t_k D s_k. \end{aligned}$$

Если $g_k = \mathbb{O}$, то x_k — решение задачи (5). Вычисления прекращаются. В противном случае находим

$$\begin{aligned} b_k &= \frac{\langle g_k, g_k \rangle}{\langle g_{k-1}, g_{k-1} \rangle}, \\ s_{k+1} &= -g_k + b_k s_k. \end{aligned}$$

По крайней мере, при $k = n$ (а возможно и раньше) получим $x_k = x_*$.

На рис. 2 схематично представлен шаг метода сопряжённых градиентов.

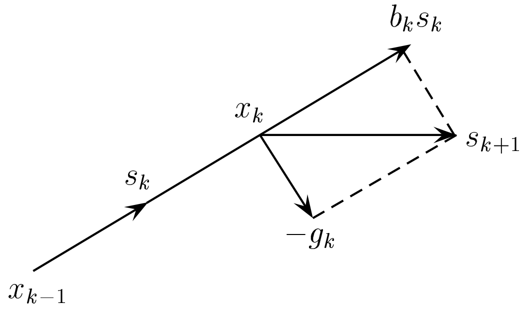

Рис. 2

5°. До сих пор мы предполагали, что матрица D в задаче (5) симметрична и положительно определена. Ослабим это предположение. Будем считать, что матрица D симметрична и неотрицательно определена. В этом случае квадратичная функция $Q(x)$ остаётся выпуклой на $\mathbb{R}^n$ и критерий оптимальности для задачи (5) сохраняет вид $Q'(x) = \mathbb{O}$.

ПРЕДЛОЖЕНИЕ 3. *Если квадратичная функция $Q(x)$ ограничена снизу на $\mathbb{R}^n$, то задача (5) имеет решение.*

Доказательство. Допустим противное. Тогда условие $Q'(x) = \mathbb{O}$ не выполняется ни при каком $x \in \mathbb{R}^n$, то есть система линейных уравнений $Dx = -c$ несовместна. Это в свою очередь означает, что найдется вектор $u_0 \in \mathbb{R}^n$, такой, что

$$Du_0 = \mathbb{O}, \quad \langle c, u_0 \rangle \neq 0.$$

Возьмём произвольный вектор $x_0 \in \mathbb{R}^n$ и запишем разложение

$$\begin{aligned} Q(x_0 + tu_0) &= Q(x_0) + t\langle Dx_0 + c, u_0 \rangle + \tfrac{1}{2}\langle Du_0, u_0 \rangle = \\ &= Q(x_0) + t\big[\langle x_0, Du_0 \rangle + \langle c, u_0 \rangle\big] = Q(x_0) + t\langle c, u_0 \rangle. \end{aligned} \tag{16}$$

Отсюда следует, вопреки условию предложения, что квадратичная функция $Q(x)$ неограничена снизу на прямой $x = x_0 + tu_0$, $t \in \mathbb{R}$.

Предложение доказано. □

Введём обозначение

$$\mathcal{P} = \{p \in \mathbb{R}^n \mid Dp = \mathbb{O}\}.$$

По-прежнему считаем, что квадратичная функция $Q(x)$ ограничена снизу на $\mathbb{R}^n$.

ПРЕДЛОЖЕНИЕ 4. *При всех $x \in \mathbb{R}^n$ и всех $p \in \mathcal{P}$ справедливо равенство*

$$\langle Q'(x), p \rangle = 0. \tag{17}$$

Д о к а з а т е л ь с т в о. Согласно (16) при $p \in \mathcal{P}$

$$Q(x_0 + tp) = Q(x_0) + t\langle c, p\rangle.$$

Учитывая ограниченность снизу функции $Q(x)$ на $\mathbb{R}^n$, заключаем, что

$$\langle c, p\rangle = 0 \quad \forall p \in \mathcal{P}.$$

Теперь имеем

$$\langle Q'(x), p\rangle = \langle Dx + c, p\rangle = \langle Dx, p\rangle = \langle x, Dp\rangle = 0.$$

Предложение доказано □

Ортогональное дополнение к линейному множеству $\mathcal{P}$ обозначим $\mathcal{P}^\perp$. Формула (17) равносильна включению

$$Q'(x) \in \mathcal{P}^\perp \quad \forall x \in \mathbb{R}^n. \tag{18}$$

ПРЕДЛОЖЕНИЕ 5. *При всех $x \in \mathcal{P}^\perp$, $x \neq \mathbb{O}$, выполняется неравенство*

$$\langle Dx, x\rangle > 0.$$

Д о к а з а т е л ь с т в о. Допустим, вопреки утверждению, что существует точка $x_0 \in \mathcal{P}^\perp$, $x_0 \neq \mathbb{O}$, в которой $\langle Dx_0, x_0\rangle = 0$. Этот факт можно проинтерпретировать следующим образом: точка x_0 доставляет минимум квадратичной функции $\varphi(x) = \frac{1}{2}\langle Dx, x\rangle$ на $\mathbb{R}^n$. Но тогда $\varphi'(x_0) = \mathbb{O}$, так что $Dx_0 = \mathbb{O}$. Получили, что $x_0 \in \mathcal{P}$. Вместе с включением $x_0 \in \mathcal{P}^\perp$ это гарантирует равенство $x_0 = \mathbb{O}$, противоречащее условию $x_0 \neq \mathbb{O}$.

Предложение доказано. □

6°. Метод сопряжённых градиентов в вырожденном случае.

Вернёмся к экстремальной задаче (5) при ослабленном условии на матрицу D. Будем считать, что матрица D симметрична и неотрицательно определена, причём $\operatorname{rank} D = r < n$. Вначале рассмотрим случай, когда квадратичная функция $Q(x)$ ограничена снизу на $\mathbb{R}^n$. Для решения задачи (5) формально воспользуемся методом сопряжённых градиентов. Вычисления будут продолжаться, пока градиенты $g_k = Q'(x_k)$ отличны от нуля. Согласно (18), $g_k \in \mathcal{P}^\perp$.

Так как сопряжённые направления s_k являются линейными комбинациями попарно ортогональных градиентов $g_0, g_1, \ldots, g_{k-1}$ и коэффициент при g_{k-1} в такой линейной комбинации равен -1, то $s_k \in \mathcal{P}^\perp$ и $s_k \neq \mathbb{O}$. В силу предложения 5, $\langle Ds_k, s_k\rangle > 0$. Это гарантирует беспрепятственную реализацию вычислений, пока $g_k \neq \mathbb{O}$.

По условию $\operatorname{rank} D = r$, так что $\dim \mathcal{P} = n - r$ и $\dim \mathcal{P}^{\perp} = r$. Текущие градиенты g_k попарно ортогональны и принадлежат $\mathcal{P}^{\perp}$. Значит, по крайней мере, при $k = r$ (а возможно и раньше) получим $g_k = \mathbb{O}$. Согласно критерию оптимальности, соответствующее x_k является решением задачи (5).

Установлено замечательное свойство метода сопряжённых градиентов: *в случае, когда* $\operatorname{rank} D = r < n$ *и квадратичная функция* $Q(x)$ *ограничена снизу на* $\mathbb{R}^n$*, метод сопряжённых градиентов решает задачу* (5) *не более, чем за* r *итераций.*

Теперь предположим, что квадратичная функция $Q(x)$ неограничена снизу на $\mathbb{R}^n$. Тогда задача (5) не имеет решения и $Q'(x) \neq \mathbb{O}$ при всех $x \in \mathbb{R}^n$. В процессе вычислений по методу сопряжённых градиентов при некотором k выполнится равенство $\langle Ds_k, s_k \rangle = 0$ (иначе процесс построения попарно ортогональных градиентов будет бесконечным). В этом случае квадратичная функция $Q(x)$ неограничена снизу на луче $x = x_{k-1} + ts_k$, $t > 0$, что следует из разложения

$$\begin{aligned} Q(x_{k-1} + ts_k) &= Q(x_{k-1}) + t\langle g_{k-1}, s_k \rangle = \\ &= Q(x_{k-1}) + t\langle g_{k-1}, -g_{k-1} + b_{k-1}s_{k-1} \rangle = Q(x_{k-1}) - t\|g_{k-1}\|^2. \end{aligned}$$

ЛИТЕРАТУРА

1. Hestenes M. R., Stiefel E. *Methods of conjudate gradients for solving linear systems* // J. Res. Nat. Bur. Standarts. 1952. Vol. 49. No. 6. P. 409–436.

2. Фаддеев Д. К., Фаддеева В. Н. *Вычислительные методы линейной алгебры.* М.: Физматгиз, 1960. 656 с.