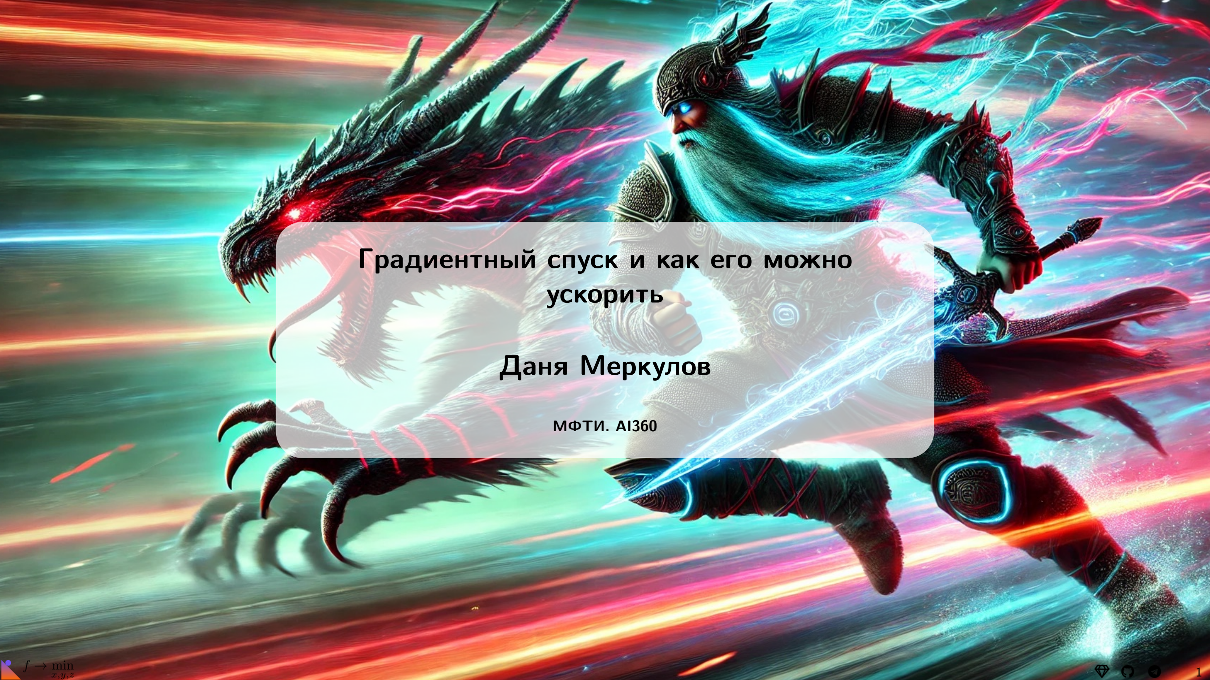


# Градиентный спуск и как его можно ускорить


**Даня Меркулов**

МФТИ. AI360

# Градиентный спуск

# Направление локального наискорейшего спуска

Рассмотрим линейное приближение дифференцируемой функции $f$ вдоль некоторого направления $h, \|h\|_2 = 1$:

# Направление локального наискорейшего спуска

Рассмотрим линейное приближение дифференцируемой функции $f$ вдоль некоторого направления $h, \|h\|_2 = 1$:

$$f(x + \alpha h) = f(x) + \alpha \langle f'(x), h \rangle + o(\alpha)$$

# Направление локального наискорейшего спуска

Рассмотрим линейное приближение дифференцируемой функции $f$ вдоль некоторого направления $h, \|h\|_2 = 1$:

$$f(x + \alpha h) = f(x) + \alpha \langle f'(x), h \rangle + o(\alpha)$$

Мы хотим, чтобы направление $h$ было направлением убывания функции:

$$f(x + \alpha h) < f(x)$$

$$f(x) + \alpha \langle f'(x), h \rangle + o(\alpha) < f(x)$$

# Направление локального наискорейшего спуска

Рассмотрим линейное приближение дифференцируемой функции $f$ вдоль некоторого направления $h, \|h\|_2 = 1$:

$$f(x + \alpha h) = f(x) + \alpha \langle f'(x), h \rangle + o(\alpha)$$

Мы хотим, чтобы направление $h$ было направлением убывания функции:

$$f(x + \alpha h) < f(x)$$

$$f(x) + \alpha \langle f'(x), h \rangle + o(\alpha) < f(x)$$

перейдя к пределу при $\alpha \to 0$:

$$\langle f'(x), h \rangle \leq 0$$

# Направление локального наискорейшего спуска

Рассмотрим линейное приближение дифференцируемой функции $f$ вдоль некоторого направления $h, \|h\|_2 = 1$:

$$f(x + \alpha h) = f(x) + \alpha \langle f'(x), h \rangle + o(\alpha)$$

Мы хотим, чтобы направление $h$ было направлением убывания функции:

$$f(x + \alpha h) < f(x)$$

$$f(x) + \alpha \langle f'(x), h \rangle + o(\alpha) < f(x)$$

перейдя к пределу при $\alpha \to 0$:

$$\langle f'(x), h \rangle \leq 0$$

Также из неравенства Коши-Буняковского-Шварца:

$$|\langle f'(x), h \rangle| \leq \|f'(x)\|_2 \|h\|_2$$
$$\langle f'(x), h \rangle \geq -\|f'(x)\|_2 \|h\|_2 = -\|f'(x)\|_2$$

# Направление локального наискорейшего спуска

Рассмотрим линейное приближение дифференцируемой функции $f$ вдоль некоторого направления $h, \|h\|_2 = 1$:

$$f(x + \alpha h) = f(x) + \alpha \langle f'(x), h \rangle + o(\alpha)$$

Мы хотим, чтобы направление $h$ было направлением убывания функции:

$$f(x + \alpha h) < f(x)$$

$$f(x) + \alpha \langle f'(x), h \rangle + o(\alpha) < f(x)$$

перейдя к пределу при $\alpha \to 0$:

$$\langle f'(x), h \rangle \leq 0$$

Также из неравенства Коши-Буняковского-Шварца:

$$|\langle f'(x), h \rangle| \leq \|f'(x)\|_2 \|h\|_2$$
$$\langle f'(x), h \rangle \geq -\|f'(x)\|_2 \|h\|_2 = -\|f'(x)\|_2$$

Таким образом, направление антиградиента

$$h = -\frac{f'(x)}{\|f'(x)\|_2}$$

даёт направление **наискорейшего локального** убывания функции $f$.

# Направление локального наискорейшего спуска

Рассмотрим линейное приближение дифференцируемой функции $f$ вдоль некоторого направления $h, \|h\|_2 = 1$:

$$f(x + \alpha h) = f(x) + \alpha \langle f'(x), h \rangle + o(\alpha)$$

Мы хотим, чтобы направление $h$ было направлением убывания функции:

$$f(x + \alpha h) < f(x)$$

$$f(x) + \alpha \langle f'(x), h \rangle + o(\alpha) < f(x)$$

перейдя к пределу при $\alpha \to 0$:

$$\langle f'(x), h \rangle \leq 0$$

Также из неравенства Коши-Буняковского-Шварца:

$$|\langle f'(x), h \rangle| \leq \|f'(x)\|_2 \|h\|_2$$
$$\langle f'(x), h \rangle \geq -\|f'(x)\|_2 \|h\|_2 = -\|f'(x)\|_2$$

Таким образом, направление антиградиента

$$h = -\frac{f'(x)}{\|f'(x)\|_2}$$

даёт направление **наискорейшего локального** убывания функции $f$.
Результатом этого является метод градиентного спуска:

$$x_{k+1} = x_k - \alpha f'(x_k)$$

# Направление локального наискорейшего спуска

Рассмотрим линейное приближение дифференцируемой функции $f$ вдоль некоторого направления $h, \|h\|_2 = 1$:

$$f(x + \alpha h) = f(x) + \alpha \langle f'(x), h \rangle + o(\alpha)$$

Мы хотим, чтобы направление $h$ было направлением убывания функции:

$$f(x + \alpha h) < f(x)$$

$$f(x) + \alpha \langle f'(x), h \rangle + o(\alpha) < f(x)$$

перейдя к пределу при $\alpha \to 0$:

$$\langle f'(x), h \rangle \leq 0$$

Также из неравенства Коши-Буняковского-Шварца:

$$|\langle f'(x), h \rangle| \leq \|f'(x)\|_2 \|h\|_2$$
$$\langle f'(x), h \rangle \geq -\|f'(x)\|_2 \|h\|_2 = -\|f'(x)\|_2$$

Таким образом, направление антиградиента

$$h = -\frac{f'(x)}{\|f'(x)\|_2}$$

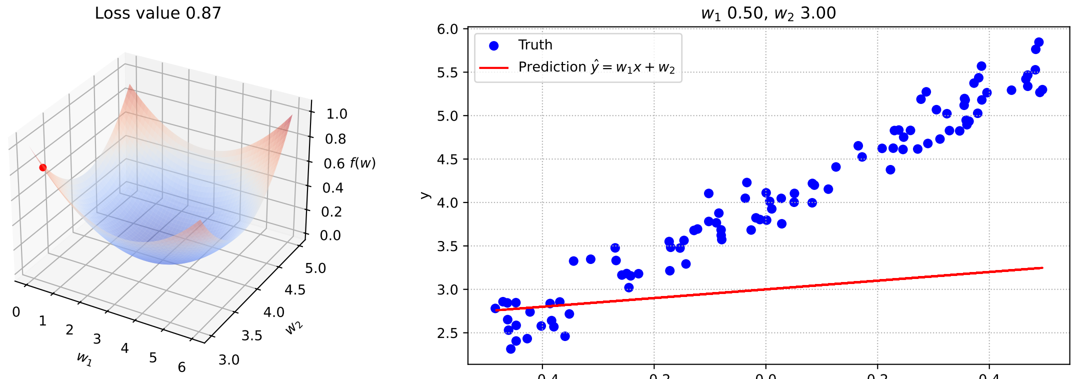
даёт направление **наискорейшего локального** убывания функции $f$.

Результатом этого является метод градиентного спуска:

$$x_{k+1} = x_k - \alpha f'(x_k)$$

# Сходимость градиентного спуска

Сходимость градиентного спуска сильно зависит от выбора шага $\alpha$:



# Наискорейший спуск

$$\alpha_k = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_{k+1}) = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_k - \alpha \nabla f(x_k))$$

Более теоретический, чем практический подход. Он также позволяет анализировать сходимость, но часто точный поиск вдоль направления может быть сложным, если вычисление функции занимает слишком много времени или стоит дорого.
Интересное теоретическое свойство этого метода заключается в том, что каждая следующая итерация ортогональна предыдущей:

$$\alpha_k = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_k - \alpha \nabla f(x_k))$$

# Наискорейший спуск

$$
\alpha_k = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_{k+1}) = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_k - \alpha \nabla f(x_k))
$$

Более теоретический, чем практический подход. Он также позволяет анализировать сходимость, но часто точный поиск вдоль направления может быть сложным, если вычисление функции занимает слишком много времени или стоит дорого.

Интересное теоретическое свойство этого метода заключается в том, что каждая следующая итерация ортогональна предыдущей:

$$
\alpha_k = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_k - \alpha \nabla f(x_k))
$$

Условие оптимальности:

# Наискорейший спуск

$$
\alpha_k = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_{k+1}) = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_k - \alpha \nabla f(x_k))
$$

Более теоретический, чем практический подход. Он также позволяет анализировать сходимость, но часто точный поиск вдоль направления может быть сложным, если вычисление функции занимает слишком много времени или стоит дорого.

Интересное теоретическое свойство этого метода заключается в том, что каждая следующая итерация ортогональна предыдущей:

$$
\alpha_k = \text{arg}\min_{\alpha \in \mathbb{R}^+} f(x_k - \alpha \nabla f(x_k))
$$

Условие оптимальности:

$$
\nabla f(x_{k+1})^\top \nabla f(x_k) = 0
$$

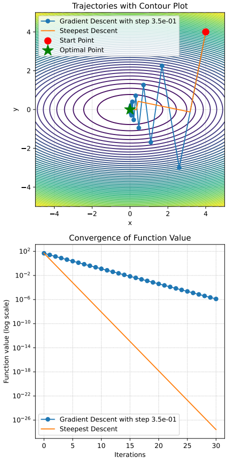



Рис. 1: Наискорейший спуск

Open In Colab ♣

# Сходимость для сильно выпуклых квадратичных функций

# Сдвиг координат

Рассмотрим следующую квадратичную оптимизационную задачу:

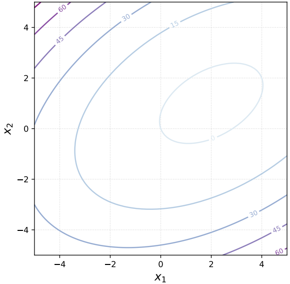
$$\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.$$

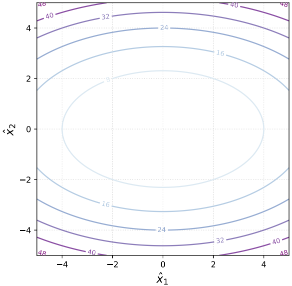
# Сдвиг координат

Рассмотрим следующую квадратичную оптимизационную задачу:

$$\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.$$

- Без ограничения общности можно положить $c = 0$, что не повлияет на процесс оптимизации.



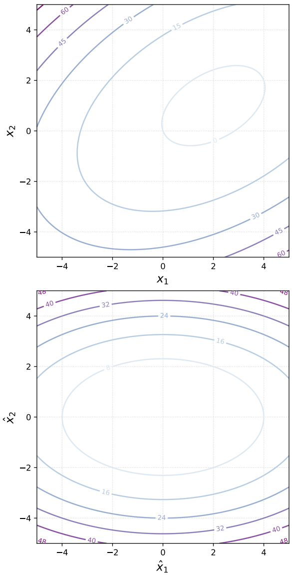



# Сдвиг координат

Рассмотрим следующую квадратичную оптимизационную задачу:

$$\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.$$

- Без ограничения общности можно положить $c = 0$, что не повлияет на процесс оптимизации.
- Второй шаг: представим матрицу $A$ в виде $A = Q \Lambda Q^T$.

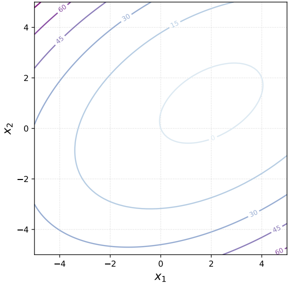


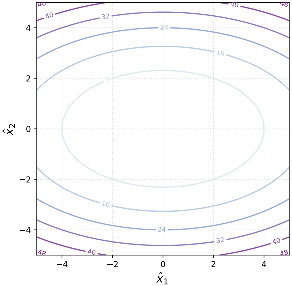
# Сдвиг координат

Рассмотрим следующую квадратичную оптимизационную задачу:

$$\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.$$

- Без ограничения общности можно положить $c = 0$, что не повлияет на процесс оптимизации.
- Второй шаг: представим матрицу $A$ в виде $A = Q \Lambda Q^T$.
- Покажем, что мы можем сделать замену координат, чтобы сделать анализ немного проще. Пусть $\hat{x} = Q^T(x - x^*)$, где $x^*$ - точка минимума исходной функции, определяемая как $Ax^* = b$. При этом $x = Q\hat{x} + x^*$.



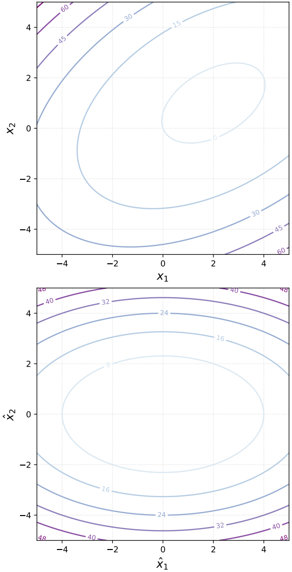
# Сдвиг координат

Рассмотрим следующую квадратичную оптимизационную задачу:

$$
\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.
$$

- Без ограничения общности можно положить $c = 0$, что не повлияет на процесс оптимизации.
- Второй шаг: представим матрицу $A$ в виде $A = Q \Lambda Q^T$.
- Покажем, что мы можем сделать замену координат, чтобы сделать анализ немного проще. Пусть $\hat{x} = Q^T(x - x^*)$, где $x^*$ - точка минимума исходной функции, определяемая как $Ax^* = b$. При этом $x = Q\hat{x} + x^*$.

$$
f(\hat{x}) = \frac{1}{2}(Q\hat{x} + x^*)^\top A (Q\hat{x} + x^*) - b^\top (Q\hat{x} + x^*)
$$



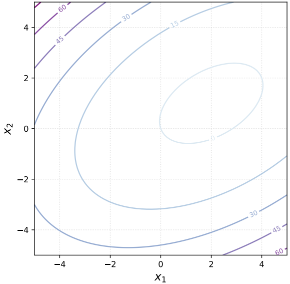
# Сдвиг координат

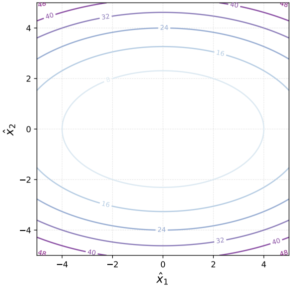
Рассмотрим следующую квадратичную оптимизационную задачу:

$$\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.$$

- Без ограничения общности можно положить $c = 0$, что не повлияет на процесс оптимизации.
- Второй шаг: представим матрицу $A$ в виде $A = Q\Lambda Q^T$.
- Покажем, что мы можем сделать замену координат, чтобы сделать анализ немного проще. Пусть $\hat{x} = Q^T(x - x^*)$, где $x^*$ - точка минимума исходной функции, определяемая как $Ax^* = b$. При этом $x = Q\hat{x} + x^*$.

$$\begin{aligned} f(\hat{x}) &= \frac{1}{2}(Q\hat{x} + x^*)^\top A(Q\hat{x} + x^*) - b^\top(Q\hat{x} + x^*) \\ &= \frac{1}{2}\hat{x}^T Q^T A Q\hat{x} + \frac{1}{2}(x^*)^T A(x^*) + (x^*)^T A Q\hat{x} - b^T Q\hat{x} - b^T x^* \end{aligned}$$



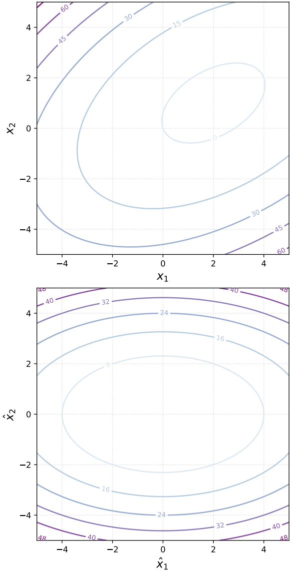
# Сдвиг координат

Рассмотрим следующую квадратичную оптимизационную задачу:

$$
\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.
$$

- Без ограничения общности можно положить $c = 0$, что не повлияет на процесс оптимизации.
- Второй шаг: представим матрицу $A$ в виде $A = Q \Lambda Q^T$.
- Покажем, что мы можем сделать замену координат, чтобы сделать анализ немного проще. Пусть $\hat{x} = Q^T(x - x^*)$, где $x^*$ - точка минимума исходной функции, определяемая как $Ax^* = b$. При этом $x = Q\hat{x} + x^*$.

$$
\begin{aligned}
f(\hat{x}) &= \frac{1}{2}(Q\hat{x} + x^*)^\top A (Q\hat{x} + x^*) - b^\top (Q\hat{x} + x^*) \\
&= \frac{1}{2}\hat{x}^T Q^T A Q \hat{x} + \frac{1}{2}(x^*)^T A (x^*) + (x^*)^T A Q \hat{x} - b^T Q \hat{x} - b^T x^* \\
&= \frac{1}{2}\hat{x}^T \Lambda \hat{x} + \frac{1}{2}(x^*)^T A (x^*) + (x^*)^T A Q \hat{x} - (x^*)^T A^T Q \hat{x} - (x^*)^T A x^*
\end{aligned}
$$



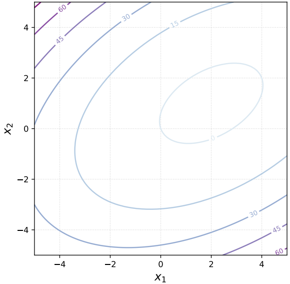
# Сдвиг координат

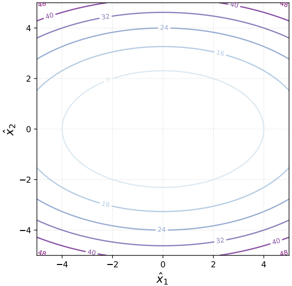
Рассмотрим следующую квадратичную оптимизационную задачу:

$$\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.$$

- Без ограничения общности можно положить $c = 0$, что не повлияет на процесс оптимизации.
- Второй шаг: представим матрицу $A$ в виде $A = Q\Lambda Q^T$.
- Покажем, что мы можем сделать замену координат, чтобы сделать анализ немного проще. Пусть $\hat{x} = Q^T(x - x^*)$, где $x^*$ - точка минимума исходной функции, определяемая как $Ax^* = b$. При этом $x = Q\hat{x} + x^*$.

$$\begin{aligned}
f(\hat{x}) &= \frac{1}{2}(Q\hat{x} + x^*)^\top A(Q\hat{x} + x^*) - b^\top(Q\hat{x} + x^*) \\
&= \frac{1}{2}\hat{x}^T Q^T A Q \hat{x} + \frac{1}{2}(x^*)^T A(x^*) + (x^*)^T A Q\hat{x} - b^T Q\hat{x} - b^T x^* \\
&= \frac{1}{2}\hat{x}^T \Lambda \hat{x} + \frac{1}{2}(x^*)^T A(x^*) + (x^*)^T A Q\hat{x} - (x^*)^T A^T Q\hat{x} - (x^*)^T A x^* \\
&= \frac{1}{2}\hat{x}^T \Lambda \hat{x} - \frac{1}{2}(x^*)^T A x^*
\end{aligned}$$



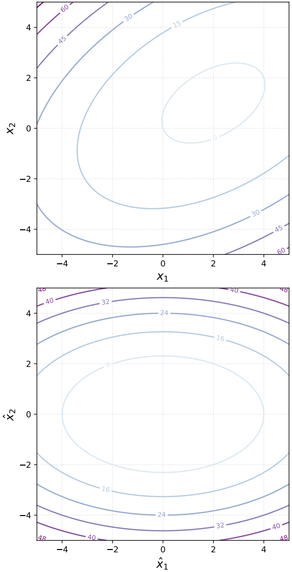
# Сдвиг координат

Рассмотрим следующую квадратичную оптимизационную задачу:

$$\min_{x \in \mathbb{R}^d} f(x) = \min_{x \in \mathbb{R}^d} \frac{1}{2} x^\top A x - b^\top x + c, \text{ where } A \in \mathbb{S}^d_{++}.$$

- Без ограничения общности можно положить $c = 0$, что не повлияет на процесс оптимизации.
- Второй шаг: представим матрицу $A$ в виде $A = Q \Lambda Q^T$.
- Покажем, что мы можем сделать замену координат, чтобы сделать анализ немного проще. Пусть $\hat{x} = Q^T(x - x^*)$, где $x^*$ - точка минимума исходной функции, определяемая как $Ax^* = b$. При этом $x = Q\hat{x} + x^*$.

$$\begin{aligned}
f(\hat{x}) &= \frac{1}{2}(Q\hat{x} + x^*)^\top A (Q\hat{x} + x^*) - b^\top (Q\hat{x} + x^*) \\
&= \frac{1}{2}\hat{x}^T Q^T A Q \hat{x} + \frac{1}{2}(x^*)^T A (x^*) + (x^*)^T A Q \hat{x} - b^T Q \hat{x} - b^T x^* \\
&= \frac{1}{2}\hat{x}^T \Lambda \hat{x} + \frac{1}{2}(x^*)^T A (x^*) + (x^*)^T A Q \hat{x} - (x^*)^T A^T Q \hat{x} - (x^*)^T A x^* \\
&= \frac{1}{2}\hat{x}^T \Lambda \hat{x} - \frac{1}{2}(x^*)^T A x^* \simeq \frac{1}{2}\hat{x}^T \Lambda \hat{x}
\end{aligned}$$



# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T \Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$x^{k+1} = x^k - \alpha^k \nabla f(x^k)$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$x^{k+1} = x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты}
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
|1 - \alpha\mu| < 1
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 \\
&-1 < 1 - \alpha\mu < 1
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 \\
&-1 < 1 - \alpha\mu < 1 \\
&\alpha < \frac{2}{\mu} \qquad \alpha\mu > 0
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$\rho(\alpha) = \max_i |1 - \alpha\lambda_{(i)}| < 1$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$\begin{aligned}
&|1 - \alpha\mu| < 1 \\
&-1 < 1 - \alpha\mu < 1 \\
&\alpha < \frac{2}{\mu} \qquad \alpha\mu > 0
\end{aligned}$$

$$|1 - \alpha L| < 1$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha\lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
|1 - \alpha\mu| &< 1 & \qquad |1 - \alpha L| &< 1 \\
-1 < 1 - \alpha\mu &< 1 & \qquad -1 < 1 - \alpha L &< 1
\end{aligned}
$$

$$
\alpha < \frac{2}{\mu} \qquad \alpha\mu > 0
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha\mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \qquad \alpha\mu > 0 && \alpha < \frac{2}{L} \qquad \alpha L > 0
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T \Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha\mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \qquad \alpha\mu > 0 && \alpha < \frac{2}{L} \qquad \alpha L > 0
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{array}{ll}
|1 - \alpha\mu| < 1 & |1 - \alpha L| < 1 \\
-1 < 1 - \alpha\mu < 1 & -1 < 1 - \alpha L < 1 \\
\alpha < \dfrac{2}{\mu} \qquad \alpha\mu > 0 & \alpha < \dfrac{2}{L} \qquad \alpha L > 0
\end{array}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha\mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \qquad \alpha\mu > 0 && \alpha < \frac{2}{L} \qquad \alpha L > 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\rho^* = \min_\alpha \rho(\alpha)
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha\lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha\mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \quad \alpha\mu > 0 && \alpha < \frac{2}{L} \quad \alpha L > 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\rho^* = \min_\alpha \rho(\alpha) = \min_\alpha \max_i |1 - \alpha\lambda_{(i)}|
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha\mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \quad\quad \alpha\mu > 0 && \alpha < \frac{2}{L} \quad\quad \alpha L > 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\begin{aligned}
\rho^* &= \min_\alpha \rho(\alpha) = \min_\alpha \max_i |1 - \alpha \lambda_{(i)}| \\
&= \min_\alpha \left\{ |1 - \alpha\mu|, |1 - \alpha L| \right\}
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T \Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha \mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha \mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \quad \alpha \mu > 0 && \alpha < \frac{2}{L} \quad \alpha L > 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\begin{aligned}
\rho^* &= \min_\alpha \rho(\alpha) = \min_\alpha \max_i |1 - \alpha \lambda_{(i)}| \\
&= \min_\alpha \{|1 - \alpha \mu|, |1 - \alpha L|\} \\
\alpha^* : &\quad 1 - \alpha^* \mu = \alpha^* L - 1
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha\mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \quad \alpha\mu > 0 && \alpha < \frac{2}{L} \quad \alpha L > 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\begin{aligned}
\rho^* &= \min_\alpha \rho(\alpha) = \min_\alpha \max_i |1 - \alpha \lambda_{(i)}| \\
&= \min_\alpha \left\{ |1 - \alpha\mu|, |1 - \alpha L| \right\} \\
\alpha^* &: \quad 1 - \alpha^*\mu = \alpha^* L - 1 \\
\alpha^* &= \frac{2}{\mu + L}
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T \Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha \mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha \mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \quad \alpha \mu > 0 && \alpha < \frac{2}{L} \quad \alpha L > 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\begin{aligned}
\rho^* &= \min_{\alpha} \rho(\alpha) = \min_{\alpha} \max_i |1 - \alpha \lambda_{(i)}| \\
&= \min_{\alpha} \left\{ |1 - \alpha \mu|, |1 - \alpha L| \right\} \\
\alpha^* &: \quad 1 - \alpha^* \mu = \alpha^* L - 1 \\
\alpha^* &= \frac{2}{\mu + L} \quad \rho^* = \frac{L - \mu}{L + \mu}
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T \Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha\mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \qquad \alpha\mu > 0 && \alpha < \frac{2}{L} \qquad \alpha L > 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\begin{aligned}
\rho^* &= \min_\alpha \rho(\alpha) = \min_\alpha \max_i |1 - \alpha \lambda_{(i)}| \\
&= \min_\alpha \left\{ |1 - \alpha\mu|, |1 - \alpha L| \right\} \\
\alpha^* &: \quad 1 - \alpha^* \mu = \alpha^* L - 1 \\
& \alpha^* = \frac{2}{\mu + L} \quad \rho^* = \frac{L - \mu}{L + \mu} \\
x^k_{(i)} &= \left( \frac{L - \mu}{L + \mu} \right)^k x^0_{(i)}
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T \Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha \lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
&|1 - \alpha\mu| < 1 && |1 - \alpha L| < 1 \\
&-1 < 1 - \alpha\mu < 1 && -1 < 1 - \alpha L < 1 \\
&\alpha < \frac{2}{\mu} \qquad \alpha\mu > 0 && \alpha < \frac{2}{L} \qquad \alpha L > 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\begin{aligned}
\rho^* &= \min_\alpha \rho(\alpha) = \min_\alpha \max_i |1 - \alpha \lambda_{(i)}| \\
&= \min_\alpha \left\{ |1 - \alpha\mu|, |1 - \alpha L| \right\} \\
\alpha^* &: \quad 1 - \alpha^*\mu = \alpha^* L - 1 \\
&\alpha^* = \frac{2}{\mu + L} \quad \rho^* = \frac{L - \mu}{L + \mu} \\
x^k_{(i)} &= \left( \frac{L - \mu}{L + \mu} \right)^k x^0_{(i)} \\
\|x^k\|_2 &\leq \left( \frac{L - \mu}{L + \mu} \right)^k \|x^0\|_2
\end{aligned}
$$

# Сходимость

Теперь мы можем работать с функцией $f(x) = \frac{1}{2}x^T\Lambda x$ с $x^* = 0$ без потери общности (убрав $\hat{x}$)

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f(x^k) = x^k - \alpha^k \Lambda x^k \\
&= (I - \alpha^k \Lambda) x^k \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)}) x^k_{(i)} \text{ Для } i\text{-й координаты} \\
x^k_{(i)} &= (1 - \alpha^k \lambda_{(i)})^k x^0_{(i)}
\end{aligned}
$$

Используем постоянный шаг $\alpha^k = \alpha$. Условие сходимости:

$$
\rho(\alpha) = \max_i |1 - \alpha\lambda_{(i)}| < 1
$$

Помним, что $\lambda_{\text{min}} = \mu > 0, \lambda_{\text{max}} = L \geq \mu$.

$$
\begin{aligned}
|1 - \alpha\mu| &< 1 & \qquad |1 - \alpha L| &< 1 \\
-1 < 1 - \alpha\mu &< 1 & \qquad -1 < 1 - \alpha L &< 1 \\
\alpha < \frac{2}{\mu} \qquad \alpha\mu &> 0 & \qquad \alpha < \frac{2}{L} \qquad \alpha L &> 0
\end{aligned}
$$

$\alpha < \frac{2}{L}$ необходимо для сходимости.

Теперь мы хотели бы настроить $\alpha$ для выбора лучшего (наименьшего) коэффициента сходимости

$$
\begin{aligned}
\rho^* &= \min_\alpha \rho(\alpha) = \min_\alpha \max_i |1 - \alpha\lambda_{(i)}| \\
&= \min_\alpha \{|1 - \alpha\mu|, |1 - \alpha L|\} \\
\alpha^* &: \quad 1 - \alpha^*\mu = \alpha^* L - 1 \\
\alpha^* &= \frac{2}{\mu + L} \quad \rho^* = \frac{L - \mu}{L + \mu} \\
x^k_{(i)} &= \left(\frac{L - \mu}{L + \mu}\right)^k x^0_{(i)} \\
\|x^k\|_2 &\leq \left(\frac{L - \mu}{L + \mu}\right)^k \|x^0\|_2 \quad f(x^k) \leq \left(\frac{L - \mu}{L + \mu}\right)^{2k} f(x^0)
\end{aligned}
$$

# Сходимость

Таким образом, мы имеем линейную сходимость в домене с коэффициентом $\frac{\varkappa-1}{\varkappa+1} = 1 - \frac{2}{\varkappa+1}$, где $\varkappa = \frac{L}{\mu}$ называется *числом обусловленности* квадратичной задачи.

| $\varkappa$ | $\rho$ | Итерации для уменьшения ошибки в 10 раз | Итерации для уменьшения невязки в 10 раз |
|---|---|---|---|
| 1.1 | 0.05 | 1 | 1 |
| 2 | 0.33 | 3 | 2 |
| 5 | 0.67 | 6 | 3 |
| 10 | 0.82 | 12 | 6 |
| 50 | 0.96 | 58 | 29 |
| 100 | 0.98 | 116 | 58 |
| 500 | 0.996 | 576 | 288 |
| 1000 | 0.998 | 1152 | 576 |

# Число обусловленности $\varkappa$