I. Постановка задачи

$$J(x) = \langle a, x\rangle^4 + 2\|x-b\|^2 \to \inf,\ x \in \mathbb{R}^5,$$

$$a = (\frac{1}{2}, \frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, \frac{3}{4}), b = (2, 4, 3, 3, 0)$$

II. Аналитическое решение

$$\langle a, b\rangle = 2, \|a\|^2 = 1, \|b\|^2 = 38$$

$$J(x) \geq 0 \Rightarrow \inf \geq 0 > -\infty$$

$$J'(x) = 3\left\langle a, x\right\rangle^3 a + 4(x-b)$$

$$J''(x)h = 12\left\langle a, x\right\rangle^2 \left\langle a, h\right\rangle a + 4h \Rightarrow$$

$$\Rightarrow \langle J(x)''h, h\rangle \geq 4\|h\|^2$$

Значит $J(x)$ - сильно выпуклая.

Так как задача без ограничений, то вариационное неравенство переходит в равенство $\Rightarrow$ $J'(x) = \theta$. Найдем отсюда $x = \alpha a + \beta b + \gamma c$, где $c \perp a, b$:

$$4\left\langle a, \alpha a + \beta b + \gamma c\right\rangle + 4(\alpha a + (\beta-1)b + \gamma c) = \theta \Rightarrow$$

$$\Rightarrow ((\alpha+2\beta)^3 + \alpha)a + (\beta-1)b + \gamma c = \theta$$

Так как a, b, c - линейно независимы, то коэффициенты при них равны нулю, в частности, $\gamma = 0$ и систему:

$$\begin{cases} \beta = 1 \\ (\alpha+2)^3 + \alpha = 0 \end{cases} \Rightarrow \begin{cases} \beta = 1 \\ \alpha = -1 \end{cases}$$

По критерию оптимальности $-a+b$ является точкой глобального минимума. Так как $J(x)$ – сильно выпуклая, то точка минимума единственная.

Ответ: $J_* = 3,\ x_* = -a + b$

III. Метод сопряженных градиентов

Приближение в данном методе строится следующим образом:

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

$$d_{k+1} = -f'(x_k) + \beta_k d_k,\ d_1 = -f'(x_1)$$

Тут настраиваемыми параметрами являются: α_k и β_k.

Будем рассматривать следующие способы выбора β_k:

1) Дай, Юань (1999 г.)

$$\beta_k^{DY} = \frac{\|f'(x_{k+1})\|^2}{\langle d_k, y_k\rangle}$$

2) Хагер, Жанг (2005 г.)

$$\beta_k^{HZ} = \left\langle y_k - 2d_k\frac{\|y_k)\|^2}{\langle d_k, y_k\rangle}, \frac{f'(x_{k+1})}{\langle d_k, y_k\rangle}\right\rangle$$

Также кроме константного выбора α_k будем еще использовать правило Армихо и искать по следующему алгоритму:

$\alpha_k \leftarrow 2 * \alpha_{k-1}$
$while\ f(x_k + \alpha_k d_k) > c * \alpha_k * \langle f'(x_k), d_k\rangle\ do$
$\quad\alpha_k \leftarrow \frac{\alpha_k}{2}$
$end\ while$
где $\alpha_0,\ c$ - это гиперпараметры.

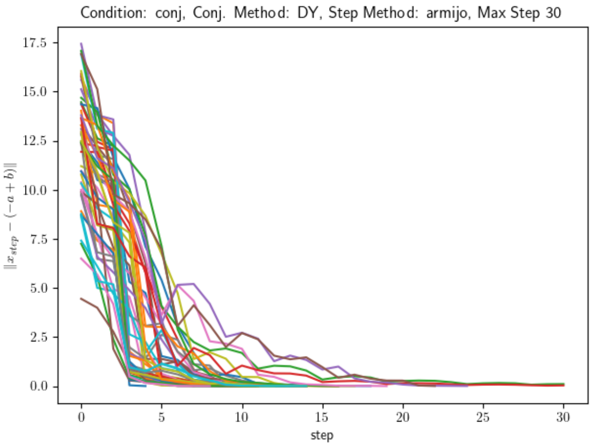
IV. Исследование сходимости от выбора методов.

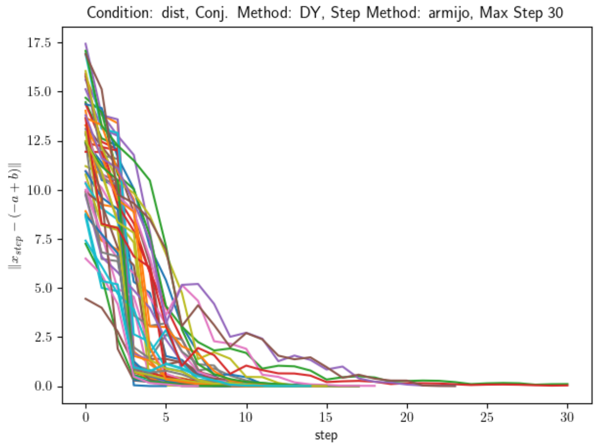
Выберем 50 произвольных точек отложенных от $-a+b$ и посмотрим зависимость расстояния точек от шага.

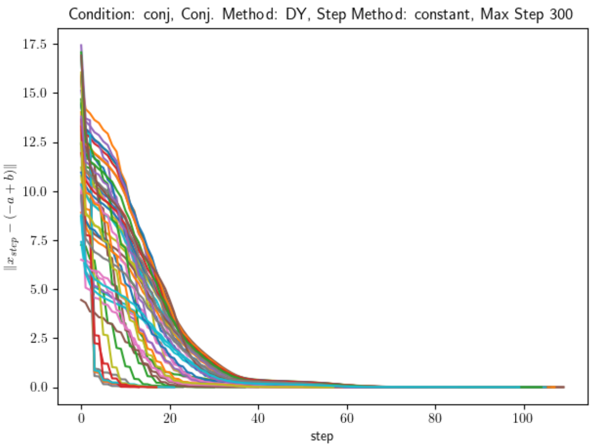
(стр.2) Ниже можно увидеть графики этой зависимости. Condition: conj (conjugate) - означает, что критерием останова является $\|d_k\| \leq \epsilon$, a dist (distance) - $\|x_k + 1 - x_k\| = \|\alpha_k d_k\| \leq \epsilon$. Зафиксируем $\epsilon = 0.01$; Для Армихо: $c = 0.25, \alpha_0 = 1$, а для константного метода: $\alpha = 0.007$. На некоторых из них можно увидеть (тяжело, но можно), что процесс при условии dist останавливается раньше как и ожидалось (т.к. α_k тоже уменьшается).

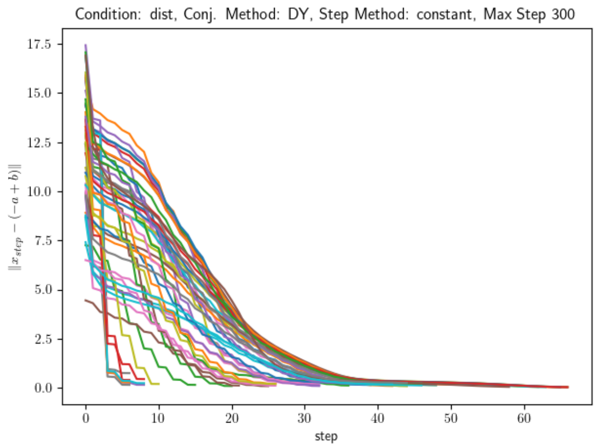
(стр.3) В случае Армихо выбирать α_0 большим или маленьким не имеет большого смысла, так как алгоритм сам это сделает, а в случае константного шага, если выбрать $\alpha = 0.5$, то получим, что результат может уйти в бесконечность и вовсе выйти за границы численного диапазона всего за 7-8 итераций. Той же проблемной ситуации можно добиться от Армихо если положить $\epsilon = 0.0000001$.

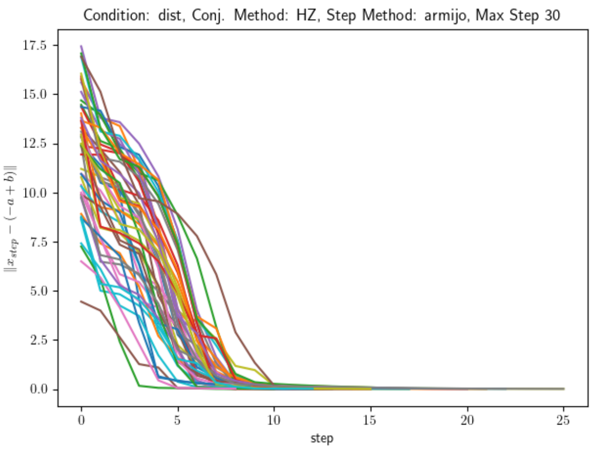
(стр.3) Можно сравнить методы DY и HZ. Для этого построим графики зависимости среднего кол-ва шагов и стандартного отклонения от расстояния начального вектора до аналитического решения при параметрах $\epsilon = 0.01, c = 0.25, \alpha_0 = 1$. На каждое расстояние (каждое второе в диапазоне от 5 до 100) было взято 500 начальных позиций.

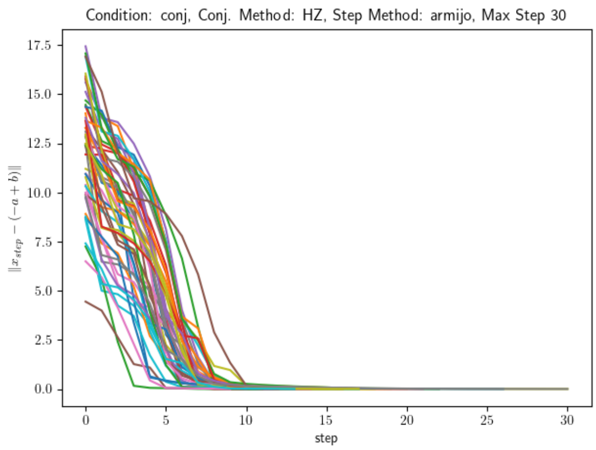

Condition: conj, Conj. Method: DY, Step Method: armijo, Max Step 30
$\|x_{step} - (-a+b)\|$
step

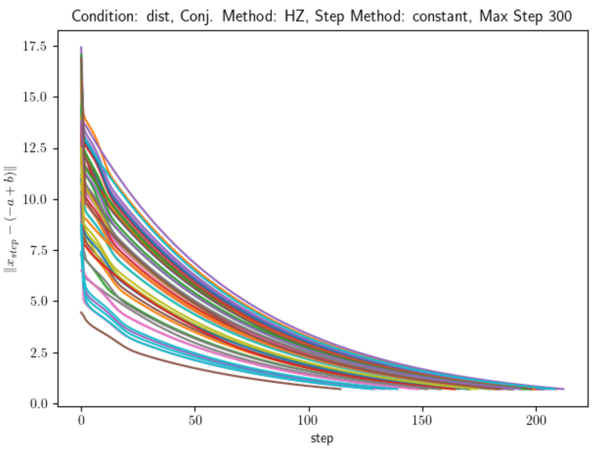

Condition: dist, Conj. Method: DY, Step Method: armijo, Max Step 30
$\|x_{step} - (-a+b)\|$
step

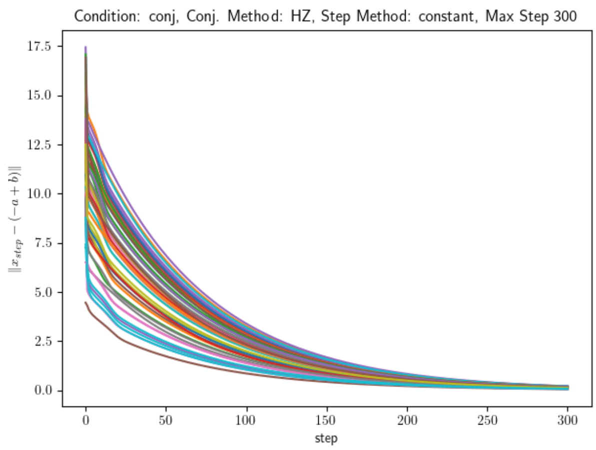

Condition: conj, Conj. Method: DY, Step Method: constant, Max Step 300
$\|x_{step} - (-a+b)\|$
step

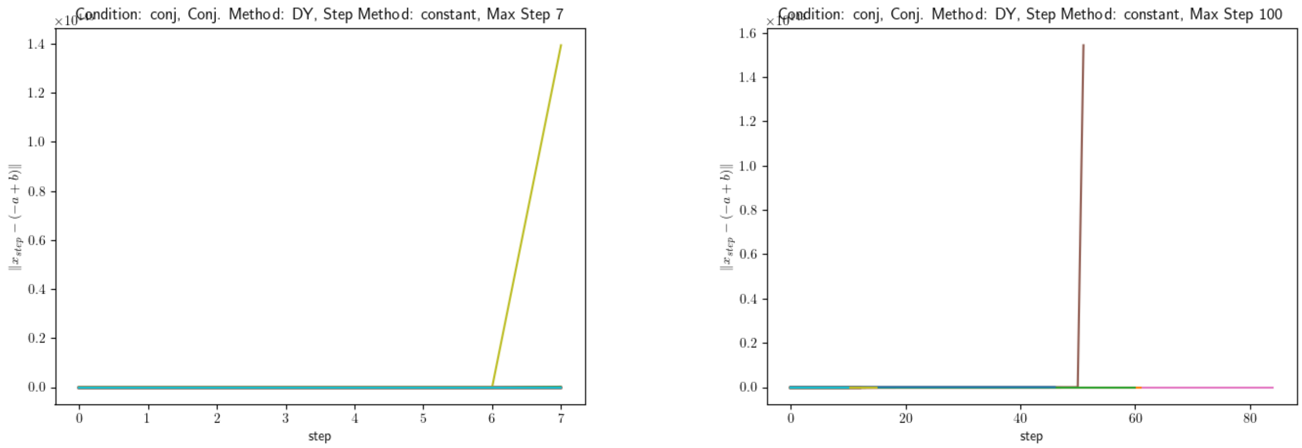

Condition: dist, Conj. Method: DY, Step Method: constant, Max Step 300
$\|x_{step} - (-a+b)\|$
step


Condition: dist, Conj. Method: HZ, Step Method: armijo, Max Step 30
$\|x_{step} - (-a+b)\|$
step

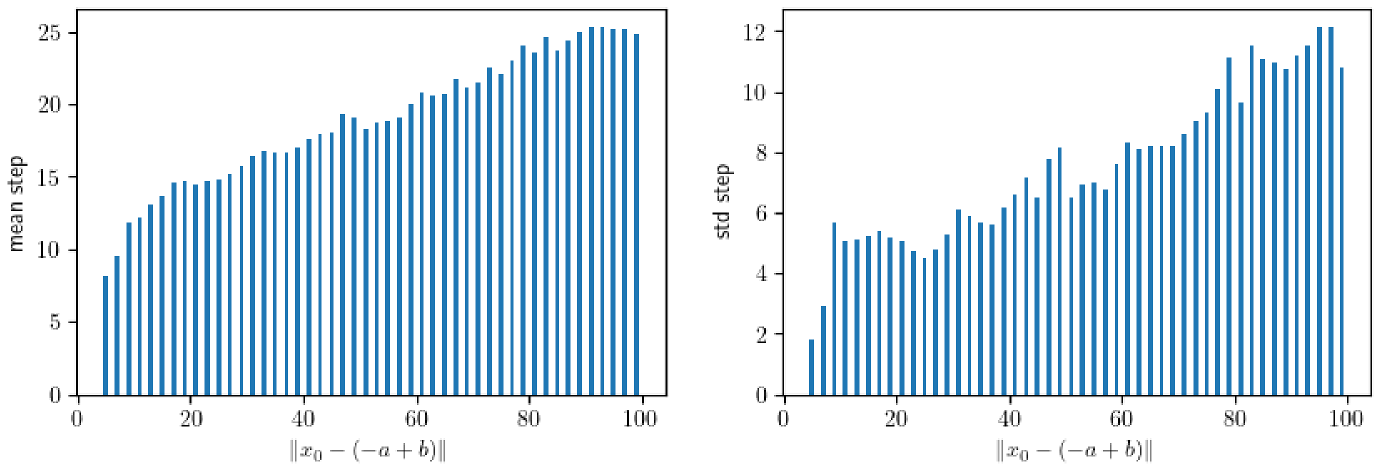

Condition: conj, Conj. Method: HZ, Step Method: armijo, Max Step 30
$\|x_{step} - (-a+b)\|$
step


Condition: dist, Conj. Method: HZ, Step Method: constant, Max Step 300
$\|x_{step} - (-a+b)\|$
step

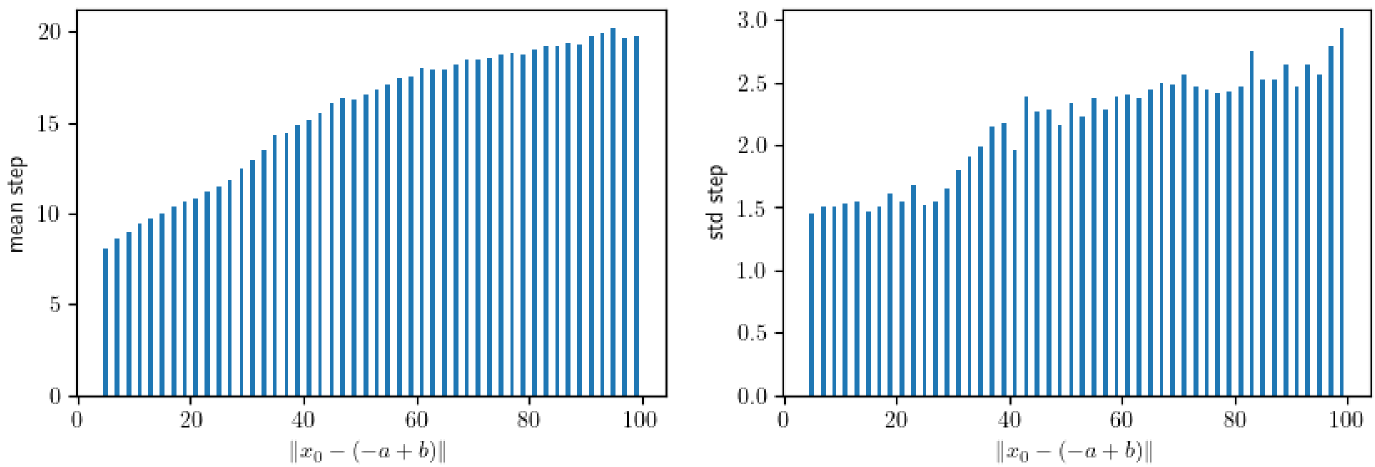

Condition: conj, Conj. Method: HZ, Step Method: constant, Max Step 300
$\|x_{step} - (-a+b)\|$
step



Condition: dist, Conj. Method: DY, Step Method: armijo, Max Step: 100



Condition: dist, Conj. Method: HZ, Step Method: armijo, Max Step: 100



V. Вывод

Хоть условие Армихо не является "самым-самым" известным сегодня, но показал себя лучше чем константный метод (конкретно при заданных параметрах и задаче). Выбор β^{HZ} более здесь предпочтителен, так как выигрывает немного в скорости, но что важнее, ведет себя более стабильно (это видно на графике стандартного отклонения).