Сириус
Международный математический центр

НЕОБХОДИМЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ И ЭКСТРЕМАЛЬНЫЕ АЛГОРИТМЫ ОПТИМИЗАЦИИ В БЕСКОНЕЧНОМЕРНЫХ ПРОСТРАНСТВАХ

Толстых Виктор Константинович
https://tolstykh.com

Донецкий государственный университет

2. Конечномерные пространства

Начнём с гладких целевых функций $J(u)$, $u \in E^n$ и далее рассмотрим дифференцируемые по Фреше функционалы $J(u)$, $u(\tau) \in L_2(S)$.
Классическое необходимое условие оптимальности (НУО):

$$\|\nabla J(u)\|_{E^n} = 0, \quad u = u_*, \tag{1}$$

Более практичным следует считать НУО:

$$\|\nabla J(u^k)\|_{E^n} \xrightarrow{k\to\infty} 0, \tag{2}$$

Эта последовательность градиентов может формироваться градиентными методами, например:

$$u^{k+1} = u^k - b^k \nabla J^k, \quad k = 0, 1, ...,$$

где шаговый множитель b^k выбирается с учётом НУО (2).

3. Особенности сходимостей

В бесконечномерном случае исчезает «приятное» свойство конечномерных пространств в виде эквивалентности сходимостей векторов по норме и по компонентам:

$$\|u^k - u_*\|_{E^n} \xrightarrow{k\to\infty} 0 \quad \Longleftrightarrow \quad \left|u_i^k - u_{*i}\right|_{\forall i=1\ldots n} \xrightarrow{k\to\infty} 0.$$

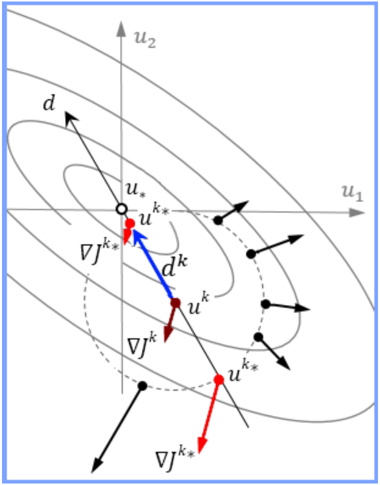
Сходимость по норме $\|u^k(\tau) - u_*(\tau)\|_{L_2(S)} \xrightarrow{k\to\infty} 0$, т.е. интегрально в среднеквадратичном, не эквивалентна равномерной сходимости $\left|u^k(\tau) - u_*(\tau)\right| \xrightarrow{k\to\infty} 0 \ \forall\tau \in S$.

Далее рассмотрим другие виды конечномерных сходимостей, которые можно было бы обобщить на бесконечномерные пространства.

4. Метод коллинеарных градиентов (МКГ)

Предлагается НУО:

$$\nabla J(u^k) \xrightarrow{k\to\infty} 0 \quad \text{коллинеарно.} \tag{3}$$



МКГ (метод первого порядка, квадратичная скорость сходимости):

$$u^{k+1} = u^k + b^k d(u^k),$$

Классическое НУО (2) игнорирует направление приближения к u_*. НУО (3) требует приближение к u_* по направлению коллинеарных градиентов, т.е. с пропорциональным уменьшением всех компонент вектора ∇J.

5. Бесконечномерные условия оптимальности

Теорема 1 (Очень сильное условие оптимальности). Если функционал $J(u)$, $u \in L_2(S)$ квадратичный и строго выпуклый, то для сходимости к оптимуму u_* на $S_\Delta \subseteq S$ необходимо и достаточно, чтобы функциональная последовательность

$$\nabla J(u^k; \tau) \xrightarrow{k \to \infty} 0 \text{ пропорционально на } S_\Delta \subseteq S. \qquad (4)$$

Следствие. Для строго выпуклого квадратичного функционала условие оптимальности (4) формирует шаг $b^0 d(u^0; \tau)$, совпадающий с шагом метода Ньютона (метод второго порядка, квадратичная скорость сходимости).

Аналог Теороемы 1 существует в конечномерном пространстве и **реализуется МКГ**. Но в бесконечномерном пространстве реализовать Теорему 1 с МКГ сложно.

6. Бесконечномерные условия оптимальности

Теорема 2 (Сильное условие оптимальности). Если строго выпуклый квадратичный функционал $J(u)$, $u \in L_2(S)$ имеет минимум в точке u_*, а последовательность управлений u^k такова, что $\left|u_* - u^k\right| \xrightarrow{k\to\infty} 0$ равномерно на S_Δ, необходимо:

$$\left|\nabla J(u^k;\tau)\right| \xrightarrow{k\to\infty} 0 \text{ равномерно на } S_\Delta \subseteq S. \qquad (5)$$

Теорема 3 — традиционная (Слабое условие оптимальности) $\|\nabla J(u^k;\tau)\|_{L_2(S)} \xrightarrow{k\to\infty} 0.$

Для строго выпуклого $J(u)$ из (5) имеем сходимость по ломаной траектории $\sum_{k=0}^{\infty} b^k \widetilde{d}^k \approx b^0 d^0$, где $b^k \widetilde{d}^k = u^{k+1} - u^k$. Близость траекторий $\sum_{k=0}^{\infty} b^k \widetilde{d}^k$ и $b^0 d^0$ определяется близостью равномерной и пропорциональной сходимостей.

7. Бесконечномерные экстремальные алгоритмы. МРНСг

Метод с регулируемым направлением спуска на основе градиента (МРНСг):

$$u^{k+1} = u^k - b^k \alpha^k \nabla J^k \text{ равномерно на } S_\Delta, \quad k = 0, 1 \ldots \qquad (6)$$

Здесь параметр регулирования направления спуска $\alpha^k(\tau) \in C^1_+(S_\Delta) = \{\alpha^k \in C^1(S_\Delta) \mid 0 < \alpha^k(\tau) < \infty \ \forall \tau \in S_\Delta\}$ задаётся из условия равномерной сходимости к оптимуму.

При каких условиях МРНСг (6) с равномерной сходимостью можно считать обобщением конечномерного МКГ с покомпонентной сходимостью?

8. Обобщение МКГ на МРНСг

МРНСг может быть обобщением конечномерного МКГ в бесконечномерное пространство, если направление минимизации МКГ

$$d^k(\tau) = -\alpha^k(\tau)\nabla J(u^k;\tau) \quad \text{— направление МКГ ,}$$

где

$$\alpha^k(\tau) = \frac{d^k(\tau)}{-\nabla J(u^k;\tau)} \qquad \text{— регулировка в МРНСг .}$$

О совпадении МРНСг и МКГ можно говорить только, если $\nabla J(u^k;\tau) \neq 0 \; \forall k, \tau, \quad \alpha^k \in C^1$, а не C^1_+. Это означает, что МРНСг в общем случае является ограниченным по сравнению с МКГ.

9. Выбор α в МРНСг для $\approx$квадратичного $J(u)$

В окрестности u^0 зададим шаблонные приближения $\tilde{u}^0(\tau)$, удовлетворяющие сильному НУО Теоремы 2:

$$u^0(\tau) \xrightarrow{\text{равномерный шаг}} \tilde{u}^0(\tau), \quad \tau \in S_\Delta,$$
$$\nabla J(u^0;\tau) \xrightarrow{\text{равномерное изменение}} \nabla J(\tilde{u}^0;\tau), \quad \tau \in S_\Delta.$$

Т.е. шаблонный шаг $\left|\tilde{u}^0 - u^0\right|$ должен быть заметным для всех $\tau \in S_\Delta$ и должен приводить к заметным на S_Δ изменениям градиента $\left|\nabla J(u^0) - \nabla J(\tilde{u}^0)\right|$. Получаем

$$\alpha(\tau) = \left|\frac{\tilde{u}^0(\tau) - u^0(\tau)}{\nabla J(u^0;\tau)}\right|, \quad \mathrm{sgn}\nabla J(u^0;\tau) = const.$$

Здесь $\mathrm{sgn} J(u^0;\tau) = const$ — это практическая реализация требования $\nabla J(u^k;\tau) \neq 0 \ \forall\tau,\ k=0$.

10. Практические рекомендации выбора α

1. МРНСг следует начинать с параметром $\alpha = 1$, т.е. с градиентного метода.
2. Шаблонный шаг «под 45°»: $\tilde{u}^0(\tau) = u^0(\tau) \pm \delta$. Получаем
$$\alpha(\tau) = \frac{\delta}{|\nabla J(u^0;\tau)|}, \quad \operatorname{sgn} J(u^0;\tau) = const, \quad \tau \in S_\Delta.$$
где $\delta > 0$ — смещение функции $u^0(\tau)$.
3. Если шаг 2 оказываются недостаточно эффективными, то следует либо сменить u^0, либо задать другую эвристическую функцию ϕ: $\tilde{u}^0(\tau) = \phi(\tau)$. При этом,
$$\alpha(\tau) = \left|\frac{\phi(\tau) - u^0(\tau)}{\nabla J(u^0;\tau)}\right|, \quad \operatorname{sgn} J(u^0;\tau) = const, \quad \tau \in S_\Delta.$$

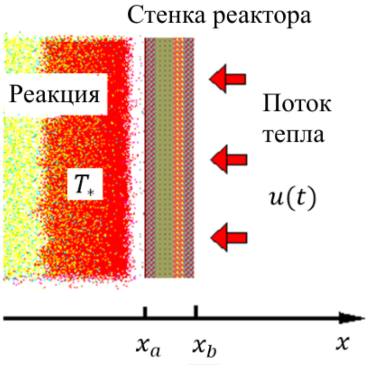
11. Иллюстрация бесконечномерной оптимизации

Одномерный нестационарный процесс теплопередачи:

$$C\rho\frac{\partial T}{\partial t} - \lambda\frac{\partial^2 T}{\partial x^2} = 0,$$

где $T(x,t)$ — температура в стенке реактора, C, ρ и λ — теплоемкость, плотность и теплопроводность.

Краевые условия:



$$\lambda\frac{\partial T}{\partial x} = q \ \text{ на } \Gamma_a = x_a \times (t_0, t_1), \qquad \lambda\frac{\partial T}{\partial x} = {\color{red}u} \ \text{ на } \Gamma_b = x_b \times (t_0, t_1),$$

$$T = T_0, \quad \text{на } [x_a, x_b] \times t_0.$$

12. Функционал и градиент

Найти плотность потока тепла $u(t) \in L_2(S)$, доставляющую минимум квадратичному целевому функционалу

$$J(u) = \int_{t_0}^{t_1} \big(T(x_a, t) - T_*(t)\big)^2 dt,$$

где $T_*(t)$ — номинальная температура реакции.

$$\nabla J(u; t) = -f \quad \text{на } S_\Delta.$$

Сопряжённая задача:

$$-C\rho \frac{\partial f}{\partial t} - \lambda \frac{\partial^2 f}{\partial x^2} = 0,$$

$$\lambda \frac{\partial f}{\partial x} = 2\big(T\big|_{\Gamma_a} - T_*\big) \quad \text{на } \Gamma_a, \qquad \lambda \frac{\partial f}{\partial x} = 0 \quad \text{на } \Gamma_b,$$

$$f = 0 \quad \text{на } [x_a, x_b] \times t_1.$$

13. Конечно-разностные аппроксимации

Расчёты состояний исходной и сопряжённой задач проводились по неявной схеме Кранка-Николсона на пространственно-временной сетке 10×100. Общее время решения задачи составляло 150 с. Выбранный шаг по времени $\Delta t = 1.5$ с оказывается достаточным для удовлетворения условию управляемости. То есть, при решении исходной задачи на отрезке времени $[t_0, t_1]$ мы будем оптимизировать управление $u(t)$ на более узком «вычислительном» множестве

$$S_\Delta = x_b \times (t_0, t_1 - \Delta t] = x_b \times [t_0 + \Delta t, t_1 - \Delta t] = x_b \times [1.5, 148.5] \text{ с.}$$

Дискретное выражение множества $S_\Delta = S$ для выбранного шага Δt.

14. Организация тестовых расчётов

Задавалось управление

$$u_*(t) = 150 + 50 \sin \frac{2\pi t}{t_1 - t_0} \, \frac{\text{кДж}}{\text{м}^2\text{с}} \quad \text{на } S,$$

рассчитывалась температура на Γ_a. Данная температура принималась номинальной (оптимальной) T_* и решалась обратная задача о восстановлении управления $u(t)$, доставляющего минимум целевому функционалу $J(u)$.

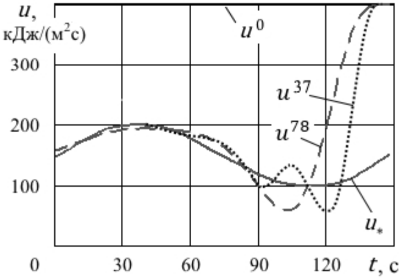
Тесты заканчивались при прекращении сходимости:

$$\frac{\|u^k - u^{k-1}\|_{L_2(S_\Delta)}}{\|u^{k-1}\|_{L_2(S_\Delta)}} \leq 10^{-6}.$$

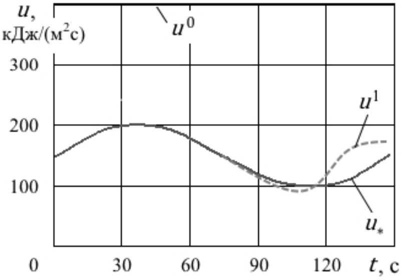
Начальное приближение

$$u^0(t) = 400 \text{ кДж/м}^2\text{с}.$$

15. Тестирование традиционных методов



(a). МНС (пунктир) и L-BFGS (точки)

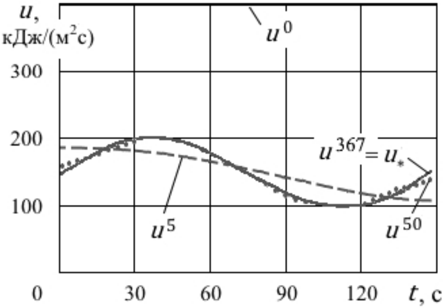


(b). Бесконечномерный МКГ

Рассматриваемый пример (квадратичность $J(u)$, вид u^0) позволил найти α при $k = 0$ для Ньютоновского шага бесконечномерного МКГ (15 подитераций МСГ). Значение $J(u)$ уменьшилось на 1 порядок лучше, чем МНС.

[При $k = 29$ уменьшилось на 4 порядка лучше, рис. почти не изменился]

16. Тестирование МРНСг



Параметр регулирования направления спуска

$$\alpha(t) = \frac{0.2u^0}{|\nabla J(u^0; t)|}.$$

Сходимость — равномерная, завершилась при $k = 367$.
Значение $J(u)$ уменьшилось на 5 порядков лучше, чем МНС.
[При $k = 50$ уменьшилось на 2 порядка лучше. Нет подитераций МКГ]

17. Выводы

Новое сильное НУО Теоремы 2, сформулированное в окрестности оптимума, открывает новые возможности конструирования бесконечномерных экстремальных алгоритмов первого порядка. Такие алгоритмы (МРНС с регулируемым направлением спуска) могут обеспечивать быструю равномерную сходимость функций-управлений к оптимуму. Традиционные бесконечномерные алгоритмы первого порядка могут реализовывать только слабое НУО со слабой (интегральной по норме) сходимостью к оптимуму.

Толстых Виктор Константинович

URL https://tolstykh.com

1. Толстых В.К. О градиенте в задачах оптимизации нестационарных систем с распределенным управлением // Вычислительные методы и программирование. 2025. Т. 26. № 3. С. 229-244.
2. Толстых В.К. Об управляемости систем с распределенными параметрами // Журн. вычисл. матем. и матем. физики, 2024, Т. 64, № 6, с. 959–972.
3. Толстых В.К. Алгоритмы оптимизации систем с многоэкстремальными функционалами // Журн. вычисл. матем. и матем. физики, 2024. Т. 64. № 3, с. 415–423.
4. Tolstykh V.K. Collinear Gradients Method for Minimizing Smooth Functions // Operations Research Forum. 2023. Vol. 4. N 20. 13 p.