Регуляризирующие алгоритмы

Факультет Космических Исследований, МГУ

Москва

Псевдорешения СЛАУ

Рассмотрим систему СЛАУ с матрицей $A \mathbb{R}^{m\times n}$, поставим задчу о псевдороешениях:

$$||Az^* - u|| = \inf\{||Az - u||_{\mathbb{R}^m} : z \in \mathbb{R}^n\}.$$

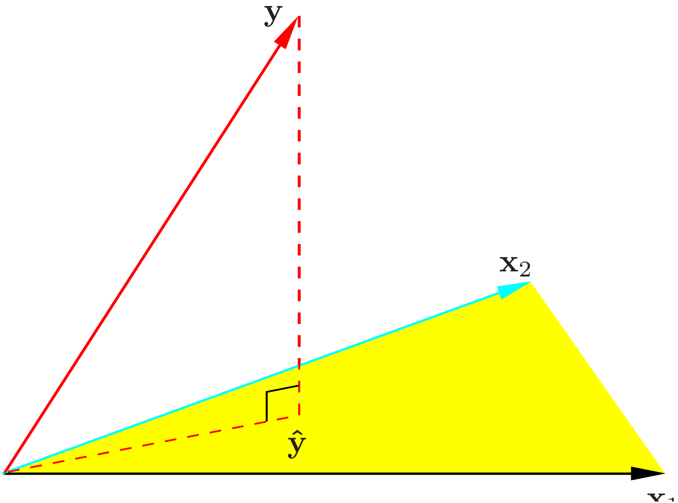
Определение

Псевдорешение $\bar{z}$ рассматривемой системы линейных уравнений, которое минимальное в норме $\mathbb{R}^m$, то есть для которых спрведливо равенство

$$||\bar{z}||_{\mathbb{R}^n} = \inf\{||z^*||_{\mathbb{R}^n} : z^* \in Z^*\},$$

называется нормальным. В случае совместной системы ее нормальное псевдорешение называется ***нормальным решением****.*

Иллюстрация проектирования



Псевдорешения СЛАУ

Определение

Оператор, ставящий в соответствие каждой правой части $u \in \mathbb{R}^m$ системы $Az = u$ ее нормальное псевдорешение, называется псевдообратным оператором, а его матрица A^+ псевдообратной матрицей для A.

В случае невырожденных квадратных матриц выполнено равенство $A^+ = A^{-1}$.

Число обусловленности

Количественная оценка СЛАУ с невырожденной матрицей можно связать с числом обусловленности матрицы

$$cond(A) = ||A||||A^{-1}||.$$

Пусть по отношению к точной системе $Az = u$ задана возмущенная $A_h z = u_\delta$, $||A - A_h|| \leq h$, $||u - u_\delta|| \leq \delta$. Возмущенная система невырождена при условии $h||A^{-1}|| < 1$, для решения возмущенной системы можно записать оценку:

$$\delta_2(z) = \frac{||z - z_\eta||}{||z||} \leq \frac{cond(A)(\delta_E(A) + \delta_2(u))}{1 - \delta_E(A) cond_E(A)}.$$

Здесь

$$cond_E(A) = ||A^{-1}||_E ||A||_E$$

- евклидово число обусловленности.

Метод сингулярного разложения

Теорема

Любую матрицу A размера $m \times n$ можно представить в виде $A = URV^T$, где – ортогональные матрицы размера $m \times m$ и $n \times n$, соответственно, а $R = diag(\rho_1, ..., \rho_M)$ - прямоугольная диагональная матрица размера $m \times n$, содержащая на диагонали неотрицательные числа $\rho_1, ..., \rho_M$, $M = \min(m, n)$, которые упорядочены по невозрастанию: $\rho_1 \geq ... \geq \rho_M \geq 0$.

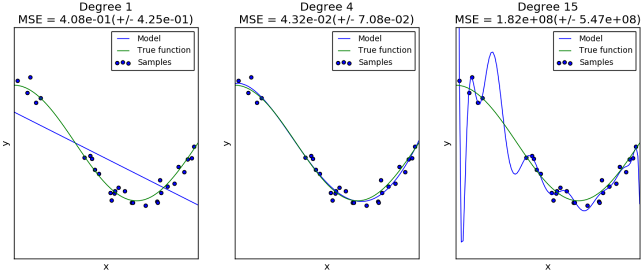
Числа ρ_k называются сингулярными числами матрицы A, при этом числа ρ_k^2, являются собственными значениями матриц AA^T, столбцы U, V - собственные вектора матриц AA^T и $A^T A$. Для матриц полного ранга можно определить спектральное число обусловленности $cond_s(A) = \rho_1 \rho_M^{-1}$.

Сингулярные числа

Сингулярные числа для операторов действующих в гильбертовых пространствах. Пусть оператор A - вполне непрерывен и не является конечномерным, то он обаладает системой сингулярных чисел $\rho_1 \geq ... \geq \rho_n \geq ... \geq 0$-собственный значения операторов A^*A, AA^*, причем $\lim_{n\to\infty} = 0$. Обратная задача с вполне непрерывным оператором A,

1. **умеренно некорректная**, если $\rho_n \asymp n^{-\nu}$ при $n \to \infty$,
2. **сильно некорректной**, если $\rho_n \asymp e^{-n\nu}$ при $n \to \infty$.

Различные виды сглаживания



Основы теории регуляризации

1. A priori информация о компактности пространства решений $\mathcal{D}$,
2. Расширение понятия решения через квазирешение

$$||Az^* - u||_U = \inf\{||Az - u||_U : z \in \mathcal{D}\} \equiv \mu,$$

3. Вариационный принцип отбора

$$\Omega[\bar{z}] = \inf\{\Omega[z] : z \in Z^*\}.$$

Понятие **регуляризирующего алгоритма**

$$z_{h\delta} = R(h, \delta, A_h, u_\delta) \to \bar{z}.$$

Возможна ли регуляризация без знания погрешности входных данных

Классические методы регуляризации используют семейство $z^\alpha = (\alpha I + A_h^* A_h)^{-1} A_h^* u_\delta = T(\alpha, A_h, u_\delta) \in Z$. При "априорном"выборе параметра $\alpha(h, \delta) \to 0$, $(h+\delta)^2/\alpha(h,\delta) \to 0$ при $h, \delta \to 0$.
Представляет интерес вопрос, можно ли сконструировать регуляризирующий алгоритм явно не зависящий от погрешности данных h, δ, $z_{h\delta} = R(A_h, u_\delta)$.
Квазиоптимальный выбор параметра: $\alpha = \alpha(A_h, u_\delta)$ находится как точка глобального минимума функции

$$\psi(\alpha) = ||\alpha \frac{dz^\alpha}{d\alpha}||^2, \ \alpha \geq 0.$$

Теорема связи шумов и непрерывности

Теорема

Пусть $\mathcal{L}$ - множество линейных непрерывных операторов, действующийх из Z в U. Пусть далее $R(A_h, u_\delta)$ - есть отображение прямого произведения $\Sigma = \mathcal{L} \times U$ в Z. Если $R(A_h, u_\delta)$ является регуляризирующим алгоритмом, не зависящим явно от h, δ, то отображение $P(A, u) = A^+ u$ определено и непрерывно на Σ.

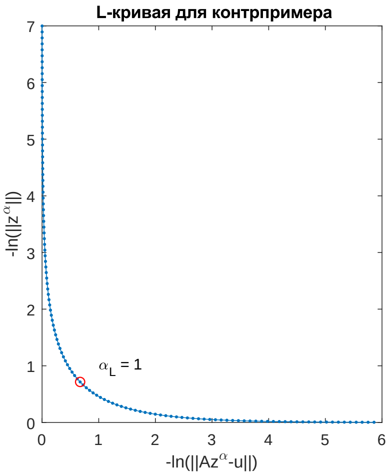
Описание метода L-кривой

Выбор параметра α, как точки максимальной кривизны линии

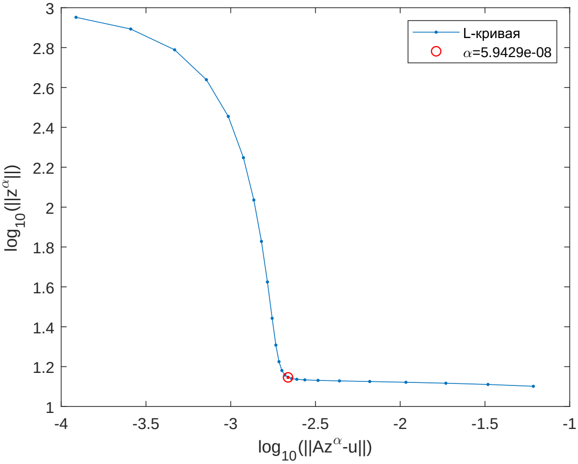
$$L = \{(ln||A_h z^\alpha - u_\delta||, ln||z^\alpha||) : \alpha > 0\}.$$

Метод подтверждался модельными расчетами без анализа ассимптотики при $h, \delta \to 0$.

Пример расчетов по методу L-кривая 1



Пример расчетов по методу L-кривая 2



Множество ограничений

1. множество невозрастающих функций $Z\downarrow_C$, ограниченных константами 0 и C,
2. множество выпуклых вверх $\check{Z}\downarrow_C$, ограниченных снизу и сверху константами 0 и C.

Первое множество компактно в $L_2[a,b]$, второе-выпуклый компакт в $L_2[a,b]$.
Численный метды регуляризации для указанных множеств могут быть основаны на градиентных методах минимизации с ограничениями, таких как условного градиента и проекции сопряженных градиентов.

Численная методы решения некорректных задач на компактных множествах

Пусть Z, U - гилбертовы пространства. Рассмотрим квадратичный функционал и его производную Фреше:

$$\Phi(z) = ||A_h z - u_\delta||^2,$$

$$\Phi'(z) = 2(A_h^* A_h z - A_h^* u_\delta).$$

Заметим $\Phi'(z)$ - удовлетворяет условию Липшица.
Для нахождения решения достаточно найти z_η, такое что $\Phi(z_\eta) \leq (\delta + h||z_\eta||)^2$ на множестве D.

Численное описание множеств

Сопоставим исследуемым множествам $Z\downarrow_C$, $\check{Z}\downarrow_C$ их конечно-разностные аппроксимации $M\downarrow_C$, $\check{M}\downarrow_C$:

$$M\downarrow_C = \{z : z \in \mathbb{R}^n,\ z_{i+1} - z_i \le 0,\ 0 \le z_i \le C,\ i = 1, .., n\},$$

$$\check{M}\downarrow_C = \{z : z \in \mathbb{R}^n, z_{i-1} - 2z_i + z_{i-1} \le 0,\ z_{i+1} - z_i \le 0,\ 0 \le z_i \le C,\ \}.$$

Регуляризация Тихонова

Рассмотрим A - линейнный взаимнооднозначный оператор действующий $Z \to U$ в гильбертовых пространствах. Действующий на D - замкнутом выпуклом множестве (известные априорные ограничения).
Необходимо построить регуляризирующий алгоритм, такой что $z_\eta = R_\eta(u_\delta, A_h) \in D$, такой что $z_\eta \to \bar{z}$ при $\eta \to 0$ $(\eta = \{\delta, h\})$.

$$M^\alpha[z] = ||Az - u||^2 + \alpha||z||^2.$$

При заданных условиях справедлива теорема о существовании и единственности минимума функционала.

Поиск минимума

1. дополнительных ограничений нет ($D = Z$)

$$(M^\alpha[z])' = 0,$$

2. дополнительные ограничений есть (или нет) - приямые методе минимизации (сопряженных градиентов, Ньютона и др.).

Уравнение Эйлера

$$A_h^* A_h z + \alpha z = A_h^* u_\delta.$$

Для случая, когда оператор A - положительно определенный и самосопряженный, пространства $Z = U$, можно применить уравнение, получаемое по методу Лаврентьева $Az + \alpha z = u_\delta$

Выбор параметра регуляризации

1. априорный способ

$$\alpha(\eta) \to 0,\ \frac{(\delta + h)^2}{\alpha(\eta)} \to 0,\ \eta \to 0,$$

   при этом доказывается $z_\eta^{\alpha(\eta)} \to \bar{z}$,

2. апостериорный способ - обобщеный принцип невязки

$$\beta_\eta(\alpha) = ||A_h z_h^\alpha - u_\delta||^2.$$

Обобщенный ПРИНЦИП невязки

Возможнц случаи, когда исходное уравнение не является разрешимым.

Определение

Мера несовместности операторного уравнения на множестве $D \in Z$ определяется

$$\mu_\eta(u_\delta, A_h) = \inf_{z \in D} ||A_h z - u_\delta||$$

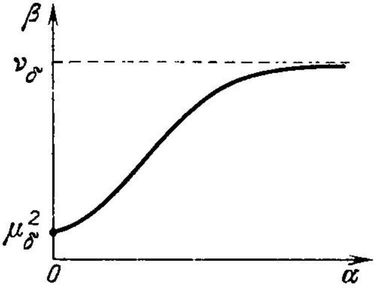
Лемма

Если $||u_\delta - \bar{u}|| \leq \delta$, $A\bar{z} = \bar{u}$, $\bar{z} \in D$, $||A_h - A|| \leq h$, то $\mu_\eta(u_\delta, A_h) \to 0$ $\eta \to 0$.

Свойства ОПН метода

Обозначим $\beta_\eta(\alpha) = ||A_h z_\eta^\alpha - u_\delta||^2$, для данной функции справедливы следующие свойства:

1. для рассматриваемого класса задач она непрерывно и монотонно не убывает,
2. $\beta_\eta(+0) \leq \mu_\eta$,
3. $\beta_\eta(+\infty) = ||u_\delta||^2$.

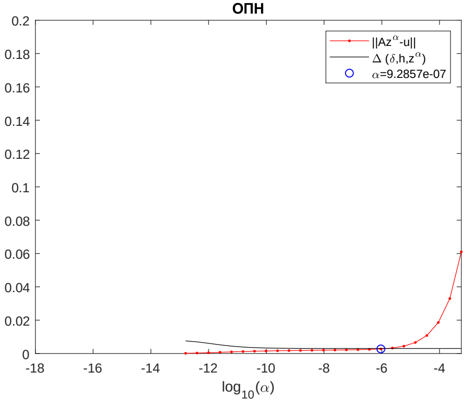


Свойства ОПН метода

Определение

Обобщенной невязкой называется функция параметра регуляризации $\alpha > 0$*:*

$$\rho_\eta(\alpha) = ||A_h z_\eta^\alpha - u_\delta||^2 - (\delta + h||z_\eta^\alpha||)^2 - (\mu_\eta(u_\delta, A_h))^2.$$



Поск решения по методу ОПН

С учетом данных свойств можно доказать, что обобщенная невязка $\rho_\eta(\alpha) = 0$ имеет единственный корень α^*, при этом соответсвующая ей экстрималь $z_\eta^{\alpha^*} \to 0$ при $\eta \to 0$. Способ нахождения решения $\rho_\eta(\alpha) = 0$:

1. уравнение Эйлера функционала Тихонова,
2. отысание корня непрерывной монотнной функции обобщенной невязки.

Обобщенный МЕТОД невязки

Рассмотрим следующую постановку
$||z_h\delta|| = \inf\{||z|| : ||A_h z - u_\delta||^2 \leq \mu_\eta + \delta + h||z||, z \in D\}$.

Теорема

Решение операторного уравнения выбранного в соответствии с ОПН эквавалентно решению экстримальной задачи ОМН.

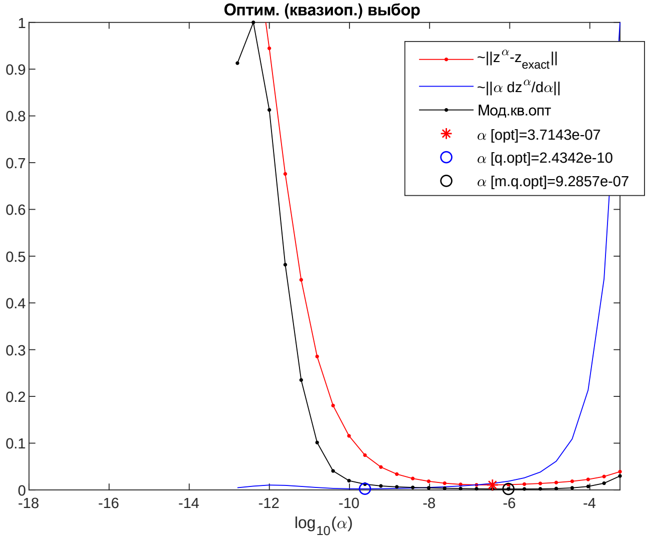
1. различный выбор нормы,...
2. поиск ближайшего по норме к заданному элементу z^0.

Модифицированный квазиоптимальный метод

Вычисляем ближайшую к нулю точку минимума так называемой модифицированной функции квазиоптимальности:

$$\Psi_{h\delta}(\alpha) = ||\alpha \frac{dz^\alpha}{d\alpha}|| + \frac{(\delta + h)^2}{\alpha}.$$

Модифицированный квазиоптимальный метод



Сравнение применения различных методов

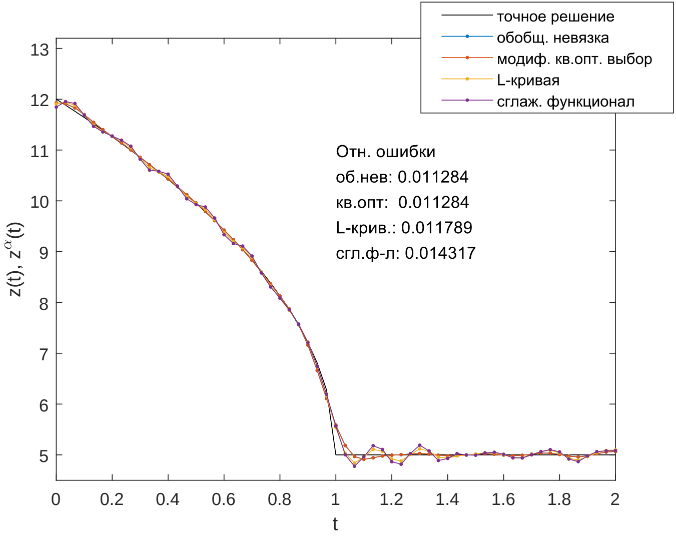
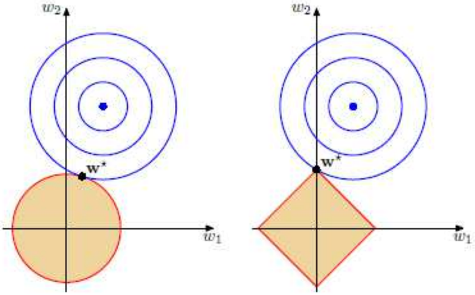
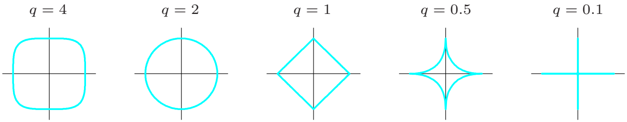
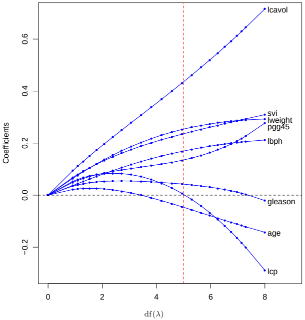


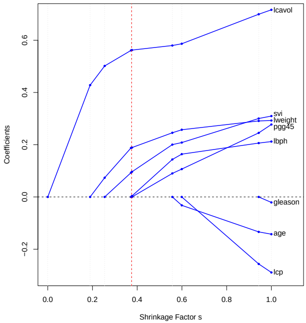
Иллюстрация различных норм





Сравнение применения L1 и L2 регуляризации





Метод сопряженных направлений

Метод минимизации квадратичных функционалов для евклидовых пространств ($x \in \mathbb{R}^n$)

$$f(x) = \frac{1}{2}(Ax, x) + (b, x) + c,$$

Требуется найти такие сопряженные направления $h^k \in \mathbb{R}^n$, $k = 0, .., n-1$, которые позволяют найти минимум $f(x)$. Пусть x^0 - произвольная начальная точка,

$$x^{k+1} = x^k + \alpha_k h^k,\; k = 0, .., n-1,$$

$$\alpha_k = \arg\min f(x^k + \alpha h^k),$$

$$f(x^n) = \min_{x \in \mathbb{R}^n} f(x).$$

Количество шагов

Теорема

Если h^k взаимно сопряженные относительно матрицы и находятся, как указано выше, то $f(x^m) = \min\limits_{x \in X_m} f(x)$, где

$$X_m = x : x = x^0 + \sum_{k=0}^{m-1} \lambda_k h^k,\ \lambda_k \in \mathbb{R}^1$$ *- афинное многообразие.*

Если $f(x)$ - сильно выпуклый функционал, то метод сопряженных градиентов может дать сверхлинейную сходимость.

Метод сопряженных градиентов

Классические формулы для построения последовательности метода сопряженных градиентов

$$x^0 \in \mathbb{R}^n,\ h^0 = -f(x^0),$$

$$h^k = -f'(x^k) + \beta_{k-1} h^{k-1},$$

$$\alpha_k = \arg\min_{\alpha \geq 0} f(x^k + \alpha h^k),$$

$$x^{k+1} = x^k + \alpha_k h^k$$

Формулы вычисления парметров метода сопряженных градиентов

По методу Флетчера-Ривза

$$\beta_{k-1} = \begin{cases} \frac{||f'(x^k)||^2}{||f'(x^{k-1})||^2}, & k = 1, n+1, 2n+1, \ldots, \\ 0, & k \neq 1, n+1, 2n+1, \ldots, \end{cases}$$

по методу Полака-Рибьера

$$\beta_{k-1} = \begin{cases} \frac{(f'(x^k), f'(x^k) - f'(x^{k-1}))}{||f'(x^{k-1})||^2}, & k \neq 1, n+1, 2n+1, \ldots, \\ 0, & k \neq 1, n+1, 2n+1, \ldots \end{cases}$$

Для квадратичного функционала возможно в явном виде получить итерационную формулу: $\alpha_k = -\frac{(f'(x^k), h^k)}{(Ah^k, h^k)}$.

Регуляризация интегрального уравнения Фредгольма первого рода

Рассмотрим оператор

$$Az = \int_a^b K(x,s)z(s)ds = u(x),$$

где ядро $K(x,s)$ - не вырождено и непрерывно в области $a \le x \le b,\ c \le s \le d$. Положим $U = L_2[c,d]$, $Z = W_2^1[a,b]$. При правильном выбора парамтра регуляризации следует сходимость в $W_2^1[a,b]$, откуда следует сходимость в $C[a,b]$.

Аппроксимация интегрального уравнения

$$M^\alpha[z] = ||A_h z - u_\delta||^2_{L_2[c,d]} + \alpha||z||^2_{W_2^1[a,b]} =$$
$$\int\limits_c^d \left[\int\limits_a^b K_h(x,s)z(s)ds - u_\delta(x)\right]^2 + \alpha\left[\int\limits_a^b z^2(s)ds + \int\limits_a^b (z'(s))^2 ds\right], \tag{1}$$

при замене $a_{ij} = K_h(x_i, s_j)$, $\hat{z} = z_{j}{}_{j=1}^{n}$, можно переписать в виде:

$$\hat{M}(\hat{z}) = \sum_{i=1}^{m}\left[\sum_{j=1}^{n} a_{ij} z_j h_s - u_i\right]^2 + \alpha\sum_{i=1}^{m} z_j^2 h_s + \alpha\sum_{i=1}^{m}(z_j - z_{j-1})^2 h_s.$$

Уравнение Эйлера Тихоновского функционала в общем виде

При использовании функционала в виде нормы пространтв $\Omega[z] = ||z||^2$

$$(\alpha L + A_h^* A_h) z = A_h^* u_\delta,$$

при дифференцируемом функционале $\Omega[z]$:

$$(\alpha \nabla \Omega[z] + A_h^* A_h) z = A_h^* u_\delta.$$

Нелинейное интегральное уравнение 1

Рассмотрим нелинейное интегральное уравнение:

$A[x, z(s)] = \int\limits_a^b K_1(x, s, z(s))ds,\ \ x \in [c, d]$, построим Тихоновский функционал

$$M^\alpha[z] = ||A_h z - u_\delta||^2 + \alpha||Lz||^2,$$

получим уравнение Эйлера

$$[(A_h)'_z] * (A_h z - u_\delta) + \alpha L^* L z = 0,$$

Нелинейное интегральное уравнение 2

производная Фреше нелиенйного оператора:

$$A'_z[v] = \int_a^b [K_1(x, s, z(s))]'_z v(s) ds,$$

так что уравнение для нахождения псевдорешения в пространстве $W_2^1[a, b]$ примет вид:

$$\int_a^b [K_1(x, s, z(s))]'_z \left(\int_a^b K_1(x, s, z(s)) ds - u_\delta(x) \right) dx + \alpha(z(t) - z''(t)) = 0$$