

Оптимальный шаг в методе Ньютона на самосогласованных функциях

Roland Hildebrand

Laboratoire Jean Kuntzmann / CNRS;
Факультет управления и прикладной математики МФТИ

Семинар Геометрическая теория оптимального управления
кафедра ОПУ мех-мата МГУ, 26 марта 2020 г.

- 1 Мотивация постановки задачи
 - Метод Ньютона и самосогласованные функции
 - Алгоритмы следования центральному пути

- 2 Анализ шага Ньютона
 - Одномерная область
 - Многомерная область

Метод Ньютона

метод Ньютона используется для поиска корня x^* векторного поля $V : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^n$ класса C^1

итерация

$$x_{k+1} = x_k - \gamma_k (V'(x_k))^{-1} V(x_k)$$

полный шаг ($\gamma_k = 1$) возвращает корень линейной модели

$$\tilde{V}_k(x) = V(x_k) + V'(x_k) \cdot (x - x_k)$$

если $\gamma_k \in (0, 1)$, то шаг *укороченный* (damped)

в качестве поля V может выступать градиент функции $F : D \rightarrow \mathbb{R}$ класса C^2

если $|V'(x^*)| \neq 0$, то метод сходится *квадратично* в некоторой окрестности x^*

Аффинная инвариантность

если поле является *ковариантным*, т.е. при замене координат $y = y(x)$ преобразуется согласно закону

$$V(x) \mapsto V(y) = \left(\frac{\partial x}{\partial y} \right)^T V(x),$$

то при *аффинных* заменах координат последовательность итераций остается неизменной

в этом случае метод Ньютона *аффинно инвариантен*

в частности, это имеет место, когда V — градиентное поле

Самосогласованные функции

вследствие аффинной инвариантности поведение метода Ньютона особенно хорошо изучается на аффинно инвариантном классе функций

для этой цели А.С. Немировский и Ю.Е. Нестеров ввели определение *самосогласованной* функции

рассмотрим функции $F : D \rightarrow \mathbb{R}$ класса C^3 , удовлетворяющие условиям $F''(x) \succ 0$ и

$$|F'''(x)[h, h, h]| \leq 2(F''(x)[h, h])^{3/2}$$

для всех $x \in D$, $h \in T_x D$

здесь $D \subset \mathbb{R}^n$ — выпуклая область

самосогласованные функции составляют замыкание по C^2 -норме множества функций, удовлетворяющих этим условиям

Самосогласованные барьеры

самосогласованные функции служат барьерами при минимизации линейных функций на выпуклых множествах \bar{D}

Определение

Самосогласованная функция $F : D \rightarrow \mathbb{R}$ называется сильно самосогласованной на области D , если $\lim_{x \in \partial D} F(x) = +\infty$.

Эллипсоиды Дикина

Теорема (Нестеров, Немировский)

Пусть функция $F : D \rightarrow \mathbb{R}$ сильно самосогласованная, и $\hat{x} \in D$.
 Тогда открытый эллипсоид Дикина

$$W(\hat{x}) = \left\{ x \mid \|x - \hat{x}\|_{\hat{x}} = \sqrt{(x - \hat{x})^T F''(\hat{x})(x - \hat{x})} < 1 \right\}$$

является подмножеством области D .

введем также эллипсоиды

$$W(\hat{x}; r) = \{x \mid \|x - \hat{x}\|_{\hat{x}} < r\}$$

радиуса r

Минимизация линейной функции

решаем задачу оптимизации с ограничениями

$$\min_{x \in D} \langle c, x \rangle$$

где $D \subset \mathbb{R}^n$ — выпуклая область

если в нашем распоряжении имеется сильно самосогласованная функция F на D , то мы можем заменить исходную задачу на задачу без ограничений

$$\min_x \tau \cdot \langle c, x \rangle + F(x),$$

где $\tau > 0$ — вещественный параметр

если $\tau \rightarrow +\infty$, то решение $x^*(\tau)$ этой задачи (если оно существует) будет стремиться к решению x^* исходной задачи

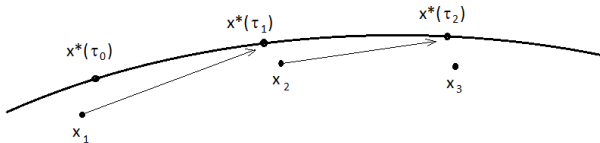
Алгоритм следования центральному пути

кривая $\sigma : \tau \mapsto x^*(\tau)$ называется *центральным путем*

задачу на поиск точки $x^*(\tau)$ решаем методом Ньютона
достаточно решить задачу приблизительно

приближенное решение x_k задачи на поиск точки $x^*(\tau_{k-1})$
используем как исходную точку для итерации Ньютона на
поиск точки $x^*(\tau_k)$

каждый шаг Ньютона чередуется с повышением параметра τ



Сходимость

условия *линейной* сходимости алгоритма

- точка x_k находится в области притяжения (квадратичной сходимости) точки $x^*(\tau_k)$
- шаг Ньютона делает прогресс, достаточный для умножения параметра τ_k на константу $\theta > 1$

для этого достаточно, если

- $\|F'(x)\|_x \leq \nu$ для всех $x \in D$ (F — самосогласованный барьер)
- $\|x^*(\tau_{k+1}) - x^*(\tau_k)\|_{x^*(\tau_k)} \geq \delta$ для всех k

тогда $\log \tau_{k+1} - \log \tau_k \gtrsim \nu^{-1} \delta$

Расстояния до текущего минимума

условие $\|x^*(\tau_{k+1}) - x^*(\tau_k)\|_{x^*(\tau_k)} \geq \delta$ выполнено, если

- x_k не дальше от $x^*(\tau_k)$, чем на $\bar{\lambda}$
- шаг Ньютона нас приближает к текущему минимуму $x^*(\tau_k)$ ближе, чем $\underline{\lambda}$

первое условие должно гарантировать второе

тогда можно отодвинуть минимум на расстояние порядка

$$\delta \sim \bar{\lambda} - \underline{\lambda}$$

Ньютоновский декремент

вернемся к минимизации функции F , с минимумом x^*
текущее положение точки x_k по отношению к минимуму
оценивается скаляром

$$\rho_k = \|F'(x_k)\|_{x_k}^* = \sqrt{F'(x_k)^T (F''(x_k))^{-1} F'(x_k)},$$

называемым *Ньютоновским декрементом*

он задает длину $\|x_{k+1} - x_k\|_{x_k}$ полного шага Ньютона
чтобы не покинуть множество D , достаточно иметь $\rho_k < 1$
чтобы сойтись, достаточно иметь $\sup \rho_{k+1} < \rho_k$

Известные оценки: полный шаг

оценки основаны на соотношении [Нестеров, Немировский]

$$(1 - \|x - x_k\|_{x_k})^2 F''(x_k) \preceq F''(x) \preceq (1 - \|x - x_k\|_{x_k})^{-2} F''(x_k)$$

для самосогласованных функций

Теорема

Пусть F — сильно самосогласованная функция на D . Если $\rho_k < \lambda^* = \frac{3-\sqrt{5}}{2} \approx 0.3820$, то после полного шага Ньютона имеем

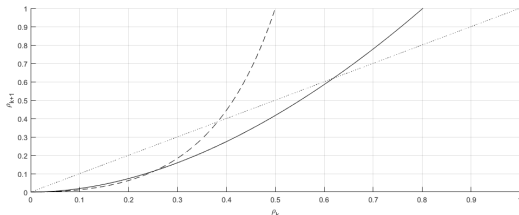
$$\rho_{k+1} \leq \left(\frac{\rho(x_k)}{1 - \rho(x_k)} \right)^2 < \rho_k.$$

Известные оценки: укороченный шаг

Теорема

Пусть F — сильно самосогласованная функция на D . Если $\rho_k < \lambda^* = \frac{\sqrt{5}-1}{2} \approx 0.6180$, то после укороченного шага Ньютона с коэффициентом $\gamma_k = \frac{1}{1+\rho_k}$ имеем

$$\rho_{k+1} \leq \frac{\rho(x_k)^2(2 + \rho(x_k))}{1 + \rho(x_k)} < \rho_k.$$



штрихованная —
полный шаг
сплошная —
укороченный шаг

Формализация задачи

анализируем шаг Ньютона с коэффициентом $\gamma \in (0, 1]$ и начальным значением декремента $\rho_k = a \in (0, 1)$ максимизируем значение декремента ρ_{k+1} в следующей точке введем переменные $g = F'$, $h = F''$, тогда динамика задается

$$g' = h, \quad h' = 2uh^{3/2}, \quad u \in U = [-1, 1]$$

декремент равен $\rho = h^{-1/2}|g|$

аффинным преобразованием достигаем

$$x_k = 0, \quad h(x_k) = 1, \quad g(x_k) = -a, \quad x_{k+1} = a\gamma$$

Задача оптимального управления

максимизируем

$$h(X)^{-1/2}|g(X)| \rightarrow \max$$

с начальными условиями

$$g(0) = -a, \quad h(0) = 1$$

и динамикой

$$g' = h, \quad h' = 2uh^{3/2}, \quad u \in U = [-1, 1]$$

на $x \in [0, X]$, $X = a\gamma$

Упрощение системы

введем новую зависимую переменную $y = h^{-1/2}g$ и независимую переменную $t = h^{1/2} \cdot (x - a\gamma)$, тогда

$$\frac{dt}{dx} = h^{1/2} \cdot (1 + ut),$$

система принимает вид

$$\dot{y} = \frac{1 - uy}{1 + ut}, \quad u \in U = [-1, 1]$$

на интервале $t \in [-a\gamma, 0]$, с начальными условиями $y(-a\gamma) = -a$ и целевой функцией

$$|y(0)| \rightarrow \max$$

про переменную h можно забыть

Принцип максимума

применяем принцип максимума Понтрягина

$$\mathcal{H} = \psi \cdot \frac{1 - uy}{1 + ut}$$

$$H = \max_{u \in U} \mathcal{H} = \begin{cases} \psi \cdot \frac{1-y}{1+t}, & \psi \cdot (t+y) < 0, \\ \psi \cdot \frac{1+y}{1-t}, & \psi \cdot (t+y) > 0. \end{cases}$$

условия трансверсальности $\psi(0) = \text{sgn } y(0)$
сопряженная переменная

$$\dot{\psi} = \begin{cases} \frac{\psi}{1+t}, & \psi \cdot (t+y) < 0, \\ -\frac{\psi}{1-t}, & \psi \cdot (t+y) > 0. \end{cases}$$

в точке $t = 0$ имеем $u = -1$

Особые режимы

задаются условием $\psi \cdot (t + y) \equiv 0$

$\psi \equiv 0$ исключено, имеем $\operatorname{sgn} \psi \equiv \operatorname{sgn} y(0)$

дифференцируем $t + y \equiv 0$:

$$1 + \frac{1 - uy}{1 + ut} = \frac{2 + u(t - y)}{1 + ut} = 2 \equiv 0$$

особых режимов нет, управление кусочно-постоянное со значениями в $\{-1, +1\}$

кривая переключения $y = -t$

Оптимальный синтез

для $y(0) > 0$ получаем решение

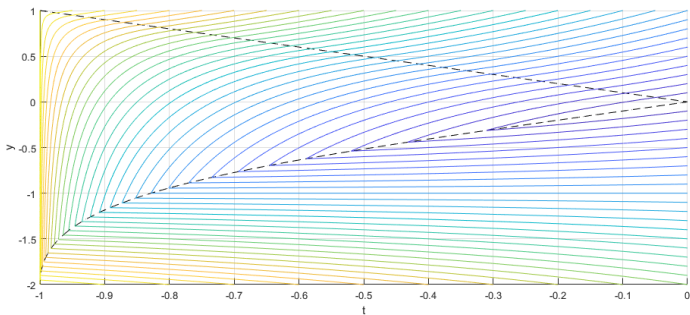
$$y(t) = \begin{cases} \frac{t+y(0)}{1-t}, & y \geq -t, \\ 1 - \frac{(2-\sqrt{1+y(0)})^2}{1+t}, & y \leq -t \end{cases}$$

для $y(0) < 0$ получаем решение

$$y(t) = \frac{t + y(0)}{1 - t}$$

траектории, исходящие из точек $\pm c$, пересекаются в точках кривой дисперсии, задаваемой соотношением

$$y = \frac{2(-1 + \sqrt{1 + t^3})}{t^2}$$



кривые дисперсии и переключения штрихованы
 между этими кривыми оптимально управление $u = +1$, иначе
 $u = -1$

Функция Беллмана

функция Беллмана постоянна на траекториях и имеет вид

$$B(t, y) = \begin{cases} -y + t + ty, & y \leq \frac{2(\sqrt{1+t^3}-1)}{t^2}, \\ (2 - \sqrt{(1-y)(1+t)})^2 - 1, & \frac{2(\sqrt{1+t^3}-1)}{t^2} \leq y \leq -t, \\ y - t - ty, & y \geq -t \end{cases}$$

подставляем начальные значения $t = -a\gamma$, $y = -a$, получаем

$$\max \rho_{k+1} = \begin{cases} a - a\gamma + a^2\gamma, & \frac{2(1-\sqrt{1-a^3\gamma^3})}{a^3\gamma^2} \leq 1, \\ (2 - \sqrt{(1+a)(1-a\gamma)})^2 - 1, & \frac{2(1-\sqrt{1-a^3\gamma^3})}{a^3\gamma^2} \geq 1 \end{cases}$$

Результат

для полного шага Ньютона ($\gamma = 1$) получаем

$$\max \rho_{k+1} = (2 - \sqrt{1 - a^2})^2 - 1$$

область сходимости $W(x_k, \lambda^*)$,

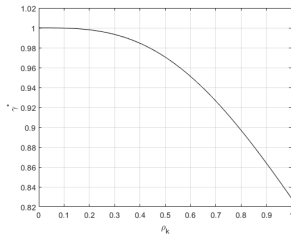
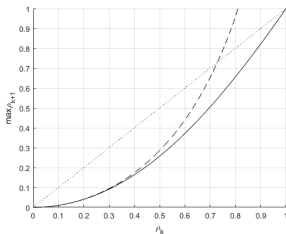
$$\lambda^* = \text{Roots}(\lambda^3 + 2\lambda^2 + 9\lambda - 8) \approx 0.7282$$

оптимальное γ^* получается минимизацией цены по γ

$$\gamma^* = \frac{2(\sqrt{1 + a^3} - 1)}{a^3}$$

$$\max \rho_{k+1} = a + \frac{2(1 - \sqrt{1 + a^3})(1 - a)}{a^2}$$

область сходимости $W(x_k)$ (весь эллипсоид Дикина)



штрихованная — полный шаг
сплошная — укороченный шаг

Множество управлений

рассмотрим общий случай функций $F : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$

если $F''(x) = I$, то допустимые значения $\frac{1}{2}F'''(x)$ — однородные кубические полиномы, ограниченные единицей на сфере S^{n-1}

определим проекцию множества \mathcal{P} таких полиномов:

$$\mathcal{U} = \{U \in \mathcal{S}^n \mid \exists p \in \mathcal{P} : U_{ij} = p_{ij1}\}$$

и множество *нижне-треугольных* матриц

$$\mathcal{V} = \left\{ V \mid \frac{V + V^T}{2} \in \mathcal{U} \right\}$$

Более простые надмножества

определим также множество симметрических матриц

$$\mathcal{U}' = \{U \in \mathcal{S}^n \mid -I \preceq U \preceq I\} \supset \mathcal{U},$$

и множество ниже-треугольных матриц

$$\mathcal{V}' = \left\{ V \mid \frac{V + V^T}{2} \in \mathcal{U}' \right\} \supset \mathcal{V}$$

в случае произвольного значения $F''(x) = WW^T$ вследствие аффинной инвариантности имеем

$$\frac{\partial^3 F}{\partial x_i \partial x_j \partial x_k} = 2 \sum_{r,s,t=1}^n W_{ir} W_{js} W_{kt} p_{rst}, \quad p \in \mathcal{P}$$

далее будем предполагать, что W — *нижне-треугольная*, тогда $W^T e_1 = W_{11} e_1$, и

$$\begin{aligned} \frac{\partial^3 F}{\partial x_i \partial x_j \partial x_1} &= 2 \sum_{r,s=1}^n W_{ir} W_{js} W_{11} p_{rs1}, \quad p \in \mathcal{P} \\ &= 2W_{11} \sum_{r,s=1}^n W_{ir} W_{js} U_{rs}, \quad U \in \mathcal{U}, \end{aligned}$$

$$F'''(x)[\cdot, \cdot, e_1] = 2W_{11} W U W^T, \quad U \in \mathcal{U}$$

Формализация задачи

введем переменные $g = F'$, $H = F'' = WW^T$

аффинным преобразованием добиваемся

$$x_k = 0, \quad g(x_k) = -ae_1, \quad H(x_k) = I, \quad x_{k+1} = a\gamma e_1$$

переходим к скалярной независимой переменной $\tau \in [0, T]$,
 $T = a\gamma$, $x(\tau) = \tau e_1$

$$\dot{g} = He_1, \quad \dot{H} = 2W_{11}WUW^T, \quad U \in \mathcal{U}$$

$$g(0) = -ae_1, \quad H(0) = I,$$

$$\sqrt{g^T(T)H^{-1}(T)g(T)} = \|W^{-1}(T)g(T)\| \rightarrow \max$$

Динамика фактора W

имеем

$$\dot{W}W^T + W\dot{W}^T = 2W_{11}WUW^T, \quad U \in \mathcal{U}$$

$$W^{-1}\dot{W} + \dot{W}^T W^{-T} = W_{11}(V + V^T), \quad V \in \mathcal{V}$$

по обе стороны стоят разложения на нижне- и
верхне-треугольную матрицу

$$\Rightarrow W^{-1}\dot{W} = W_{11}V, \quad \dot{W} = W_{11}WV, \quad V \in \mathcal{V}$$

Упрощение системы

введем новую зависимую переменную $y = W^{-1}g$ и независимую переменную $t = W_{11} \cdot (\tau - a\gamma)$, тогда

$$\frac{dt}{d\tau} = W_{11} \cdot (1 + tV_{11}),$$

система принимает вид

$$\dot{y} = \frac{e_1 - Vy}{1 + tV_{11}}, \quad V \in \mathcal{V}$$

на интервале $t \in [-a\gamma, 0]$, с начальными условиями $y(-a\gamma) = -ae_1$ и целевой функцией

$$\|y(0)\| \rightarrow \max$$

Принцип максимума

применяем принцип максимума Понтрягина

$$\mathcal{H}(t, y, p, V) = \frac{\langle p, -Vy + e_1 \rangle}{tV_{11} + 1},$$

$$H(t, y, p) = \max_{V \in \mathcal{V}} \frac{\langle p, -Vy + e_1 \rangle}{tV_{11} + 1},$$

условия трансверсальности $p(0) = \frac{y(0)}{\|y(0)\|}$

особые режимы: H независимо от V только в том случае,
когда p и y параллельны e_1

это снова одномерная ситуация, в которой особых режимов нет

Вращательная симметрия

пусть O — ортогональная матрица такая, что $Oe_1 = e_1$, тогда

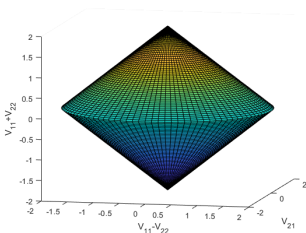
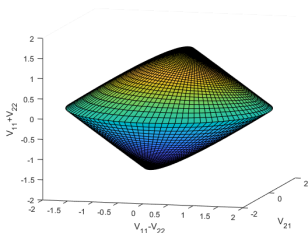
$$H(t, y, p) = H(t, Oy, Op)$$

если в момент t векторы y, p, e_1 находятся в плоскости P , то $\frac{\partial H}{\partial y}, \frac{\partial H}{\partial p} \in P$, и $y, p \in P$ для всех t

ситуация имеет место при $t = 0$

поэтому можно считать $n = 2$

Множество управлений



слева — \mathcal{V} , справа — \mathcal{V}'

окружность \mathcal{C} на границе общая и состоит из матриц вида

$$V = \begin{pmatrix} \cos \phi & 0 \\ 2 \sin \phi & -\cos \phi \end{pmatrix}, \quad \phi \in [-\pi, \pi]$$

Максимизация функции Понтрягина

максимум дробно-линейной функции на выпуклом компакте достигается в экстремальной точке

максимизируем сначала на надмножестве \mathcal{V}' :

кроме точек \mathcal{C} , \mathcal{V}' имеет экстремальные точки $V = \pm 1$

$$\hat{H} := \max_{V \in \mathcal{C}} \mathcal{H}$$

$$= \frac{p_1 + (p_1 y_1 - p_2 y_2)t + \sqrt{(p_1 y_1 - p_2 y_2 + p_1 t)^2 + 4p_2^2 y_1^2 (1 - t^2)}}{1 - t^2}$$

$$V = \pm 1 \Rightarrow \mathcal{H} = \frac{p_1 \mp (p_1 y_1 + p_2 y_2)}{1 \pm t}$$

значит $\max_{V \in \mathcal{V}'} \mathcal{H} = \max_{V \in \mathcal{C}} \mathcal{H}$, если

$$\sqrt{(p_1 y_1 - p_2 y_2 + p_1 t)^2 + 4p_2^2 y_1^2 (1 - t^2)} \geq -p_1 t - p_1 y_1 - p_2 y_2 + 2p_2 y_2 t$$

$$\sqrt{(p_1 y_1 - p_2 y_2 + p_1 t)^2 + 4p_2^2 y_1^2 (1 - t^2)} \geq p_1 t + p_1 y_1 + p_2 y_2 + 2p_2 y_2 t$$

в этом случае также $\max_{V \in \mathcal{V}'} \mathcal{H} = \max_{V \in \mathcal{C}} \mathcal{H}$ и $H = \hat{H}$

a posteriori эти неравенства выполняются строго

Промежуточное резюме

максимум декремента ρ_{k+1} задается значением нормы $\|y(0)\|$, где $(y(t), p(t))$, $t \in [-a\gamma, 0]$ — траектория гамильтоновой системы с гамильтонианом

$$\frac{p_1 + (p_1 y_1 - p_2 y_2)t + \sqrt{(p_1 y_1 - p_2 y_2 + p_1 t)^2 + 4p_2^2 y_1^2 (1 - t^2)}}{1 - t^2}$$

и краевыми условиями

$$y(-a\gamma) = -ae_1, \quad p(0) = \frac{y(0)}{\|y(0)\|}$$

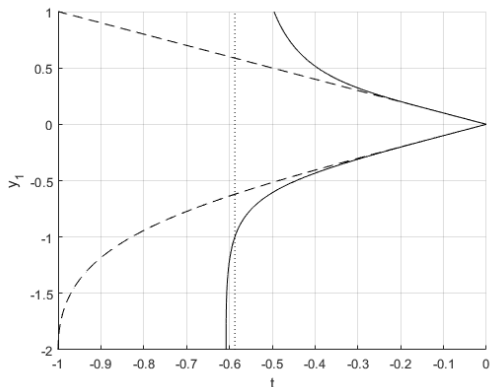
Фокальные точки

систему можно проинтегрировать с конца (динамическое программирование)

выпускаем траекторию из точки $(t, y, p) = (0, y_0, \frac{y_0}{\|y_0\|})$ для всех $y_0 \in \mathbb{R}^2$

в частности, для $y_0 = (c, 0)$ получаем траектории одномерного синтеза

линеаризовав систему вокруг такой траектории, посмотрим, где фундаментальная система решений вырождается
это определит *фокальную точку*, за которой 1D траектория перестает быть оптимальной



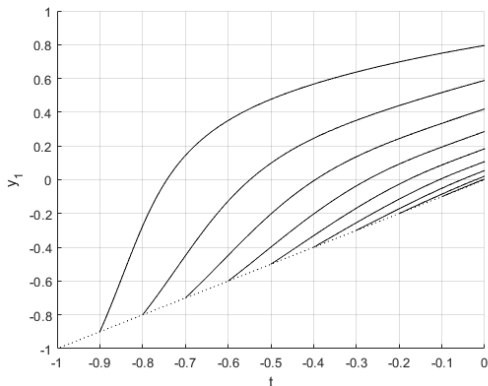
сплошная — множество фокальных точек

штрихованная — кривая переключения и дисперсионная кривая

пунктирная — асимптота кривой фокальных точек

Полный шаг Ньютона

прямая $y_1 = t$ ($\gamma = 1$) лежит за кривой фокальных точек: на оптимальной траектории y_2, p_2 не равны нулю
проекция оптимальных траекторий на (t, y_1) -плоскость



Минимизация по γ

чтобы получить оптимальную длину шага, минимизируем цену по γ при фиксированном a
начальное время будет варьировать (но это \min max, не свободный конец)

минимум достигается при $\frac{\partial B(t,y)}{\partial t} = -H = 0$

Лемма

Гиперповерхность $H = 0$ является интегральной поверхностью гамильтоновой системы.

подставляем $p(0) = \frac{y(0)}{\|y(0)\|}$ в $H(0) = 0$:
при $t = 0$ имеем $(y_1 + \frac{1}{2})^2 + y_2^2 = \frac{1}{4}$

Доказательство леммы:

$$\dot{H} = \frac{\partial H}{\partial t} =$$

$$H \cdot \frac{p_1 y_1 - p_2 y_2 + p_1 t + t \sqrt{(p_1 y_1 - p_2 y_2 + p_1 t)^2 + 4p_2^2 y_1^2 (1 - t^2)}}{\sqrt{(p_1 y_1 - p_2 y_2 + p_1 t)^2 + 4p_2^2 y_1^2 (1 - t^2)}(1 - t^2)}$$

значит $H = 0 \Rightarrow \dot{H} = 0$

Понижение размерности

предположим, что на оптимальных траекториях $y_2 \geq 0$ (если нет, применим симметрию $(y_2, p_2) \mapsto (-y_2, -p_2)$)

из $H = 0$ получим

$$(y_1^2 - 1)p_1^2 - 2y_1y_2p_1p_2 + (4y_1^2 + y_2^2)p_2^2 = 0,$$

$$\frac{p_1}{p_2} = \frac{y_1y_2 + \sqrt{-4y_1^4 + 4y_1^2 + y_2^2}}{y_1^2 - 1}$$

система примет вид

$$\dot{y}_1 = \frac{-p_1 y_1^2 + p_2 y_2 y_1 + p_1}{p_1 + p_1 y_1 t - p_2 y_2 t} =$$

$$\frac{\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}(-y_2(y_1 + t) + (y_1 t + 1))\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}}{4y_1^4 t^2 + 8y_1^3 t + 4y_1^2 - y_2^2 t^2 + y_2^2}$$

$$\dot{y}_2 = -\frac{4p_2 y_1^2 - p_1 y_1 y_2 + p_2 y_2^2}{p_1 + p_1 y_1 t - p_2 y_2 t} =$$

$$\frac{\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}(4ty_1^3 + 4y_1^2 + y_2^2 - ty_2\sqrt{-4y_1^4 + 4y_1^2 + y_2^2})}{4y_1^4 t^2 + 8y_1^3 t + 4y_1^2 - y_2^2 t^2 + y_2^2}$$

более того,

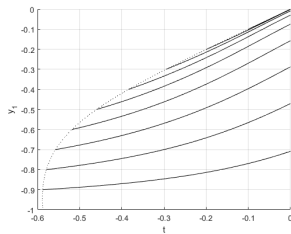
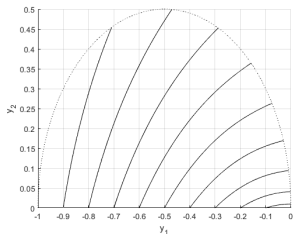
$$\frac{dy_2}{dy_1} = \frac{\sqrt{4y_1^2(1-y_1^2) + y_2^2} + y_1y_2}{1-y_1^2}$$

не зависит от t ;

$$\frac{dt}{dy_1} = \frac{y_2(y_1 + t) + (y_1t + 1)\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}}{(1-y_1^2)\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}}$$

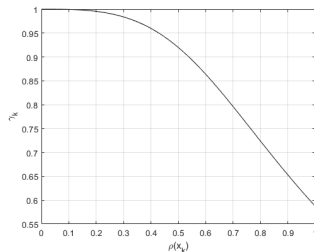
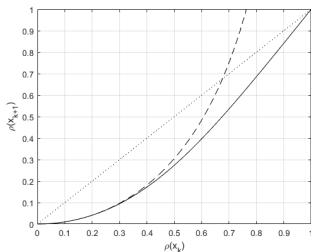
линейно по t

интегрируем первое ур-е от $y = (-a, 0)$ до полукруга $(y_1 + \frac{1}{2})^2 + y_2^2 = \frac{1}{4}$, а второе обратно по полученной траектории с начальным значением $t = 0$ на этом полукруге



проекции оптимальных траекторий

Результаты



слева: $\max \rho_{k+1}$ для полного шага (сплошная) и для оптимального шага (штрихованная)

справа: оптимальный коэффициент γ^*

Асимптотика

для оптимального шага имеем разложения

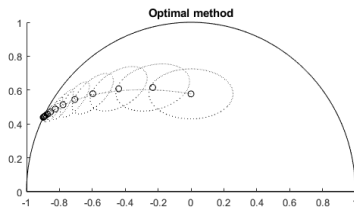
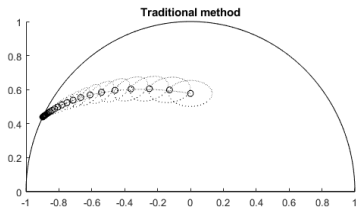
$$\max \rho_{k+1} = \rho_k^2 - \frac{1}{4} \rho_k^4 \log \rho_k + \left(\frac{\log 2}{2} - \frac{1}{16} \right) \rho_k^4 + o(\rho_k^5)$$
$$\gamma^* = 1 - \frac{\rho_k^3}{2} + \frac{\rho_k^4}{4} + O(\rho_k^5 \log \rho_k)$$

имеем также

$$\lim_{\rho_k \rightarrow 1} \gamma^* = 2^{2/3} - 1 \approx 0.5874$$

область сходимости для оптимального шага есть весь эллипсоид Дикина $W(x_k)$

Алгоритмы



слева: отодвигаем минимум настолько, чтобы иметь $\rho_k = \rho^* \approx 0.2291$ и делаем полный шаг Ньютона

справа: отодвигаем минимум настолько, чтобы иметь $\rho_k = \rho^* \approx 0.4429$ и делаем оптимальный шаг Ньютона

ρ^* в обоих случаях максимизирует $\rho_k - \max \rho_{k-1}$, т.е. минимальное гарантированное продвижение

Преимущества

точный анализ шага Ньютона позволяет в алгоритмах следования центральному пути

- расширить область сходимости вокруг текущего минимума
- увеличить длину шага вдоль центрального пути
- делать меньше итераций для достижения данной точности
- терпеть больше погрешности между текущей точкой и центральным путем

следующий этап: учесть в задаче оптимального управления направление, касательное к центральному пути

Спасибо за внимание