Лекция 13

Условная минимизация

Методы модифицированной функции Лагранжа

функция Лагранжа

$$L(\vec{x},\vec{\lambda}) = f(\vec{x}) + \sum_i^p \lambda_i\, g_i = f(\vec{x}) + \vec{\lambda}^t \vec{c}(\vec{x})$$

При определенных условиях точка $(\vec{x}^*,\vec{\lambda}^*)$ может быть седловой для этой функции, т.е.

$$L(\vec{x}^*,\vec{\lambda}) \le L(\vec{x}^*,\vec{\lambda}^*) \le L(\vec{x},\vec{\lambda}^*)$$

Но найти $\vec{x}^*$ посредством минимизация функции $L(\vec{x},\vec{\lambda}^*)$ нельзя, т.к. вектор $\vec{\lambda}^*$неизвестен. В более же общем случае, даже зная вектор $\vec{\lambda}^*$, этот поиск может не дать результата.

Методы модифицированной функции Лагранжа

модифицированная функция Лагранжа

$$L_A(\vec{x},\vec{\lambda}) = f(\vec{x}) + \vec{\lambda}^t\vec{c}(\vec{x}) + \frac{\rho}{2}\vec{c}^t(\vec{x})\vec{c}(\vec{x}),$$

где скаляр $\rho > 0$.

Всегда существует конечное значение $\hat{\rho}$ такое, что при $\rho > \hat{\rho}$ матрица Гессе $H_{L_A}(\vec{x}^*,\vec{\lambda}^*)$ будет положительно определенной и в точке $\vec{x}^*$ функция $L_A(\vec{x},\vec{\lambda}^*)$ будет иметь минимум.

Методы модифицированной функции Лагранжа

Пример

Найти $x^* = \operatorname{argmin} x^3$, при $-x - 1 \leq 0$.

Из условия Куна-Такера:

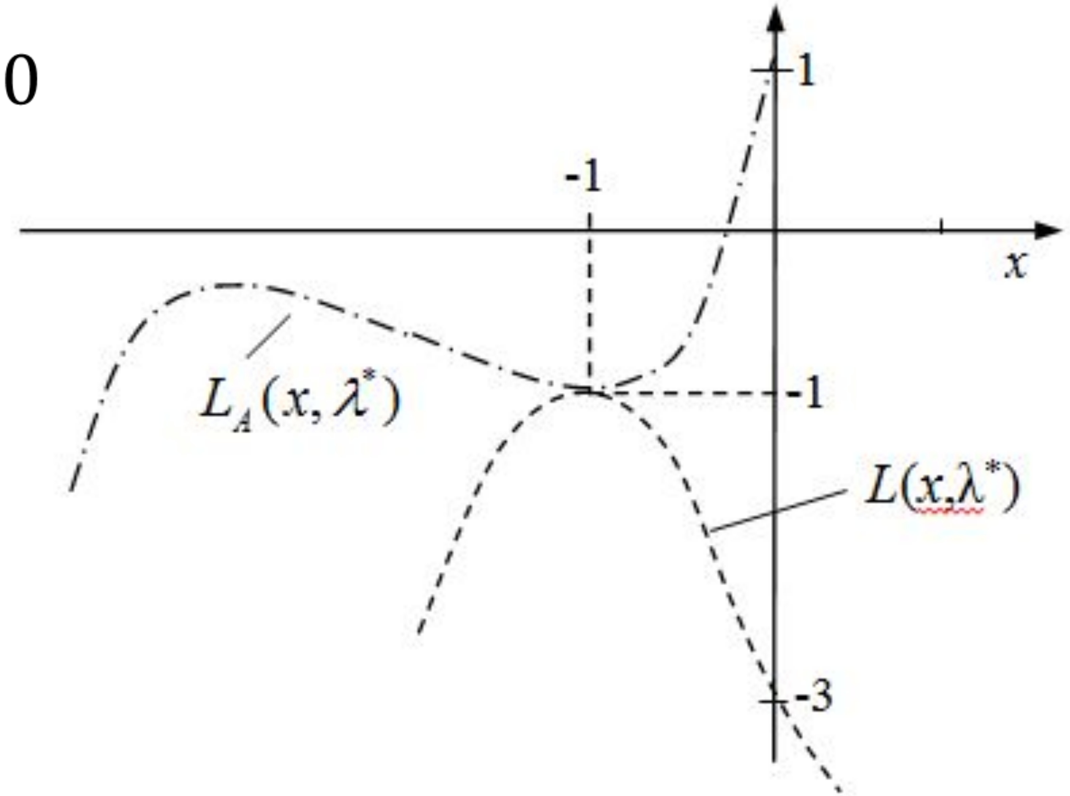
$$3x^2\Big|_{x=-1} + \lambda(-1) = 0 \Rightarrow \lambda^* = 3 \geq 0$$

Функция Лагранжа

$$L(x,\lambda^*) = x^3 - 3(x+1) \Rightarrow x^* = -1$$

Модифицированная функция Лагранжа

$$L_A(x,\lambda^*) = x^3 - 3(x+1) + \frac{\rho}{2}(x+1)^2.$$



Методы модифицированной функции Лагранжа

Алгоритм

<u>Ш0</u>. Задать $\vec{\lambda}_0, \vec{x}_0, \vec{c}_0, \rho_0$. $k = 0$.

<u>Ш1</u>. Проверить выполнение критериев останова.

<u>Ш2</u>. Найти $\vec{x}^*_{k+1} = \text{argmin} L_A(\vec{x}, \vec{\lambda}_k, \rho_k)$.

<u>Ш3</u>. Пересчитать оценки $\vec{\lambda}_{k+1}$ и уточнить состав активных ограничений $\vec{c}$.

<u>Ш4</u>. Если $\sum_{i \in I} g_i^2(\vec{x}_{k+1}) \approx \sum_{i \in I} g_i^2(\vec{x}_k)$ то $\rho := 2 \cdot \rho$

<u>Ш5</u>. $k := k + 1$. Перейти к Ш2.

Методы модифицированной функции Лагранжа

Модификация Пауэлла

$$L_A = f(\vec{x}) + \sum_{i=1}^{m}\left[< g_i(\vec{x}) + \sigma_i^{(k)} >^2 - \left(\sigma_i^{(k)}\right)^2\right] + \sum_{i=1}^{l}\left\{\left[h_i(\vec{x}) + \tau_i^{(k)}\right]^2 - (\tau_i^{(k)})^2\right\},$$

где $\sigma_i^{(k)} =< g_i\ (\vec{x}^*_{(k-1)}) + \sigma_i^{(k-1)} >$ и $\tau_i^{(k)} = h_i\ (\vec{x}^*_{(k-1)}) + \tau_i^{(k-1)}$

$\vec{x}^*_{(k-1)}$ – точка безусловного минимума функции $L_A(\vec{x})$, полученная при значениях $\vec{\sigma}_{k-1}$ и $\vec{\tau}_{k-1}$

Методы модифицированной функции Лагранжа

$$\nabla L_A(\vec{x}) = \nabla f + 2\sum_{i=1}^{m} < \mathrm{g_i}\,(\vec{x}) + \sigma_i^{(k)} > \times \nabla g_i(\vec{x}) + 2\sum_{i=1}^{l} \left[h_i(\vec{x}) + \tau_i^{(k)}\right] \times \nabla h_i(\vec{x})$$

$$H_{L_A}(\vec{x}) = \nabla^2 f + 2\sum_{i=1}^{m} < \mathrm{g_i}\,(\vec{x}) + \sigma_i^{(k)} > \nabla^2 g_i \; + [\nabla g_i(\vec{x})]^2 + 2\sum_{i=1}^{l} \left[h_i(\vec{x}) + \tau_i^{(k)}\right] \cdot \nabla^2 h_i(\vec{x}) + [\nabla h_i(\vec{x})]^2$$

Полученное выражение показывает, что в случае линейных $g_i(\vec{x})$ и $h_i(\vec{x})$ матрица H_{L_A} вообще не зависит от $\sigma_i^{(k)}$ и $\tau_i^{(k)}$ и, следовательно, топология функции L_A не изменяется с изменением этих параметров.

Методы модифицированной функции Лагранжа

Покажем, что точка $\vec{x}^* = \lim\limits_{k\to\infty} \vec{x}_k^*$ является точкой Куна-Такера.

Последовательности $\{\sigma^{(k)}\}$ и $\{\tau^{(k)}\}$ будут сходиться к своим пределам σ^* и τ^*вместе со сходимостью последовательности $\vec{x}_k^* \to \vec{x}^*$.

$$\tau_i^{k+1} = h_i(\vec{x}^{(k)}) + \tau_i^{(k)}$$

$$\tau_i^{k+1} - \tau_i^{k} = h_i(\vec{x}^{(k)})$$

$$\Rightarrow$$

$$h_i(\vec{x}^*) \underset{k\to\infty}{\to} 0$$

Методы модифицированной функции Лагранжа

$$\lim_{k\to\infty} \sigma_i^{(k)} = \sigma_i^* \qquad \lim_{k\to\infty} g_i(\vec{x}_k) = g_i(\vec{x}^*) \qquad \sigma_i^* =< g_i(\vec{x}^*) + \sigma_i{}^* >$$

для неактивных ограничений

$$g_i(\vec{x}^*) < 0 \;\Rightarrow\; < g_i(\vec{x}^*) + \sigma_i{}^* >= 0 \;\Rightarrow\; \sigma_i{}^* = 0$$

$$+$$

для активных ограничений

$$g_i(\vec{x}^*) = 0 \;\Rightarrow\; \sigma_i^* =< g_i(\vec{x}^*) + \sigma_i{}^* > \;\Rightarrow\; \sigma_i{}^* \geq 0$$

$$=$$

$$\underset{i=1,2,\ldots,m}{\sigma_i^* g_i(\vec{x}^*) = 0}$$

$$\nabla f(\vec{x}^*) + 2\sum_{i\in I} \sigma_i^* \nabla \mathrm{g_i}\,(\vec{x}^*) + 2\sum_{i=1}^{k} \tau_i^* \nabla h_i\,(\vec{x}^*) = 0.$$

Методы модифицированной функции Лагранжа

Пример

Найти условный минимум функции $f(\vec{x}) = (x_1 - 4)^2 + (x_2 - 4)^2$ при ограничении $h(x) = x_1 + x_2 - 5 = 0$

Модифицированная функция Лагранжа по версии Пауэлла:

$$L_A(\vec{x}, \tau_k) = (x_1 - 4)^2 + (x_2 - 4)^2 + (x_1 + x_2 - 5 + \tau_k)^2 - {\tau_k}^2$$

$$\frac{\partial L_A}{\partial x_1} = 2(x_1 - 4) + 2(x_1 + x_2 - 5 + \tau_k) = 0$$

$$\frac{\partial L_A}{\partial x_2} = 2(x_2 - 4) + 2(x_1 + x_2 - 5 + \tau_k) = 0.$$

Методы модифицированной функции Лагранжа

$$x_{1,k}^{*} = x_{2,k}^{*} = 3 - \frac{\tau_k}{3} \qquad \tau_{k+1} = h(\vec{x}_k^{*}) + \tau_k$$

0.0	(3.0, 3.0)	3.0	1.0
1	(2.667, 2.667)	4.3333	0.33
1.333	(2.556, 2.556)	4.482	0.11
1.4453	(2.5185, 2.5185)	4,498	0.037
1.48183	(2.5062, 2.5062)	4,499	0.012
1,5	(2.5, 2.5)	4.5	0.0

Методы модифицированной функции Лагранжа