II. Условная оптимизация

6 марта 2019 г. 11:46

Условная минимизация

- задачи поиска экстремума функции на множестве, заданном ограничениями в виде равенств и/или неравенств

$\overline{1,m} \coloneqq 1, \dots, m$

Задача нелинейного программирования

Задача минимизации функции f на множестве х, заданном набором ограничений
$X = \{\vec{x} \in \mathbb{R}^N, g_i(\vec{x}) \leq 0; i = \overline{1,m}\}$
$g_1, \dots, g_m \in C^1$ ($\nabla g_1, \dots, \nabla g_m -$ существуют и непрерывны на X)

Ограничения типа равенства

Пусть $X = \{\vec{x} \in \mathbb{R}^2, g(x_1, x_2) = 0\}$. Нужно найти минимум $f \in C^1$ на X.
Пусть $\forall x_1, x_2$: $\frac{\partial g(x_1,x_2)}{\partial x_1} \neq 0 \Rightarrow g(x_1, x_2) = 0$ **разрешимо относительно $\boldsymbol{x}$**,
т.е. $\forall x_1, x_2$: $g(x_1, x_2) = 0 \Rightarrow x_1 = \gamma(x_2); \min_X f = \min_{x_2} f(\gamma(x_2), x_2)$

$f, \gamma -$ дифференцируемы. Условие экстремальности:

$$\frac{\partial f}{\partial x_1} \cdot \frac{\partial \gamma}{\partial x_2} + \frac{\partial f}{\partial x_2} = 0$$

$$g(\gamma(x_2), x_2) \equiv 0 \Rightarrow \frac{\partial g}{\partial x_1} \cdot \frac{\partial \gamma}{\partial x_2} + \frac{\partial g}{\partial x_2} = 0 \Rightarrow \frac{\partial \gamma}{\partial x_2} = -\frac{\partial g}{\partial x_2} \cdot \left(\frac{\partial g}{\partial x_1}\right)^{-1}$$

С учётом условия экстремальности:

$$-\frac{\partial f}{\partial x_1} \cdot \frac{\partial g}{\partial x_2} \cdot \left(\frac{\partial g}{\partial x_1}\right)^{-1} + \frac{\partial f}{\partial x_2} = 0$$

Обозначим $\lambda \coloneqq \frac{\partial f}{\partial x_1} \cdot \left(\frac{\partial g}{\partial x_1}\right)^{-1} \Bigg|_{x_1^*, x_2^*}$. Тогда

$$\begin{cases} \dfrac{\partial f}{\partial x_2} + \lambda \dfrac{\partial g}{\partial x_2} = 0 \\ \dfrac{\partial f}{\partial x_1} + \lambda \dfrac{\partial g}{\partial x_1} = 0 \\ g(x_1, x_2) = 0 \end{cases}$$

В точке минимума выполняются эти 3 соотношения. Их можно записать через функцию Лагранжа:
$F(\vec{x}, \lambda) = f + \gamma g$
Тогда необходимое условие минимума может быть записано следующим образом:

$$\begin{cases} \dfrac{\partial F}{\partial x_2} = \dfrac{\partial f}{\partial x_2} + \lambda \dfrac{\partial g}{\partial x_2} = 0 \\ \dfrac{\partial F}{\partial x_1} = \dfrac{\partial f}{\partial x_1} + \lambda \dfrac{\partial g}{\partial x_1} = 0 \\ \quad \dfrac{\partial F}{\partial \lambda} = g(x_1, x_2) \end{cases}$$

Таким образом, задача условной минимизации сведена к задаче безусловной минимизации.

Ограничения типа неравенств

Задача поиска минимума f на $X = \{\vec{x} \in \mathbb{R}^n, g_i(\vec{x}) \leq 0; i = \overline{1, m}\}$, где f, g_i имеют непрерывные производные по всем аргументам в некотором открытом подмножестве, содержащем множество X.

Задача - получить необходимые условия экстремума f на X.

Пусть $\vec{x}^*$ — экстремальная точка. Свяжем с $\vec{x}^*$ множество индексов активных ограничений $I(\vec{x}^*) = \{i : g_i(\vec{x}^*) = 0; i \in \{1 \dots m\}\}$.

Лемма. Пусть $\exists \vec{S}, \forall i \in I(\vec{x}^*)$: $\begin{cases} (\nabla g_i(\vec{x}^*), S) < 0 \\ (\nabla f(x^*), S) < 0 \end{cases}$

Тогда $\vec{x}^*$ не является точкой локального минимума.

Доказательство. Пусть $\varepsilon > 0$. Тогда с помощью ряда Тейлора:

$g_i(\vec{x}^* + \varepsilon\vec{S}) = g_i(\vec{x}^*) + \varepsilon(\nabla g_i(\vec{x}^*), S) + o(\varepsilon)$

$f(\vec{x}^* + \varepsilon\vec{S}) = f(\vec{x}^*) + \varepsilon(\nabla f(\vec{x}^*), \vec{S}) + o(\varepsilon)$

Пусть $i \in I(\vec{x}^*) \Rightarrow g_i(\vec{x}^*) = 0, g_i(\vec{x}^* + \varepsilon\vec{S}) < 0$.

Если $i \notin I(\vec{x}^*)$, то $g_i(\vec{x}^*) < 0 \Rightarrow$ при всех i и достаточно малых положительных ε: $g_i(\vec{x}^* + \varepsilon\vec{S}) < 0$. Таким образом, при достаточно малых $\varepsilon > 0$ точка $\vec{x}^* + \varepsilon\vec{S}$ допустима и функция f на этом луче убывает. Значит, $\vec{x}^*$ — не экстремальна ■

Лемма Фаркаша

$\forall A_{m \times n}$ - матрицы $m \times n$ справедливо ровно одно из следующих условий:

- $\exists \vec{x} \in R^n : A\vec{x} < 0$ — все кординаты вектора $A\vec{x}$ отрицательны
- $\exists \vec{\lambda} \in \mathbb{R}^m, A^T\vec{\lambda} = 0, \lambda_i \geq 0, i \coloneqq \overline{1, m}$

Без доказательства

Теорема Каруша-Джона

Пусть $\vec{x}^*$ — экстремальная точка задачи нелинейного программирования. Пусть в точке $\vec{x}^*$ градиенты функций, соответствующие активным ограничениям, линейно независимы.

Тогда $\exists \lambda_1, \dots, \lambda_m \geq 0$:

$$f(x) = \begin{cases} \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0 \\ \lambda_i g_i(x^*) = 0, i = 1, \dots, m \end{cases}$$

Доказательство

$$\nexists S: \begin{cases} \left(\nabla g_i(\vec{x}^*), \vec{S}\right) < 0 \\ \left(\nabla f(\vec{x}^*), \vec{S}\right) < 0 \end{cases} \forall i \in I(\vec{x}^*)$$

Составим матрицу:

$$A = \begin{pmatrix} \nabla g_i(\vec{x}^*) \\ \dots \\ \nabla f(\vec{x}^*) \end{pmatrix}, i \in I(\vec{x}^*)$$

$\exists \vec{S}: A\vec{S} < 0$. Значит, по лемме Фаркаша $\exists \lambda_0, \dots, \lambda_m \geq 0$:

$$\lambda_0 \nabla f(\vec{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\vec{x}^*) = 0$$

(полагаем $\lambda_i = 0$, если $i \notin I(\vec{x}^*)$)

Для активных ограничений $g_i(\vec{x}^*) = 0$, для неактивных $\lambda_i = 0$. Тогда $\lambda_i g_i(\vec{x}^*) = 0; i = \overline{1, m}$.
$\lambda_0 \neq 0$, т.к. иначе градиенты соответствующих активных ограничений были бы линейно зависимы.
Поделив предыдущее выражение на λ_0, получим требуемое утверждение.

Пример
Найти минимум $f(x_1, x_2) = x_1 x_2$ при ограничении $x_1^2 + x_2^2 \leq 1$.

$f(x_1, x_2) = x_1 x_2; g(x) = x_1^2 + x_2^2 \leq 1$

$$\begin{cases} \dfrac{\partial f}{\partial x_1} = x_2 \\ \dfrac{\partial f}{\partial x_2} = x_1 \end{cases} \Rightarrow \nabla f(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$g(x) = x_1^2 + x_2^2 - 1 \leq 0$$

$$\begin{cases} \dfrac{\partial g}{\partial x_1} = 2x_1 \\ \dfrac{\partial g}{\partial x_2} = 2x_2 \end{cases} \Rightarrow \nabla g = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

$$\Rightarrow \begin{cases} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \lambda \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = 0 \\ \lambda(x_1^2 + x_2^2 - 1) = 0 \end{cases}$$

$$\Rightarrow \begin{cases} x_2 + 2\lambda x_1 = 0 \\ x_1 + 2\lambda x_2 = 0 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases} \Rightarrow$$

$$\Rightarrow \left[\begin{array}{l} \begin{cases} \lambda = -\dfrac{1}{2} \\ x_1 = x_2 = \pm\dfrac{1}{\sqrt{2}} \end{cases} \\ \begin{cases} \lambda = \dfrac{1}{2} \\ x_1 = \pm\dfrac{1}{\sqrt{2}} \\ x_2 = -x_1 \end{cases} \end{array} \right.$$

Задача выпуклого программирования

Пусть $g_1, \dots, g_m, f$ — выпуклы. Задача - поиск минимума f на $X = \{\vec{x} \in \mathbb{R}^n, g_i(\vec{x}) \leq 0, i = 1, \dots, m\}$.

Утверждение. Допустимое множество в задаче выпуклого программирования выпукло.

Доказательство. Пусть $\vec{x}_1, \vec{x}_2 \in X, \lambda \in [0,1]$.
Рассмотрим $\vec{z} = \lambda\overrightarrow{x_1} + (1-\lambda)x_2 \in X$. Т.к .$\mathbb{R}^n$ — выпукло, то $z \in \mathbb{R}^n$.
Надо проверить $g_i(\vec{z}) < 0$. По свойству выпуклости g:
$g_i(\lambda\vec{x}_1 + (1-\lambda)\vec{x}_1) \leq \lambda g_i(\vec{x}_1) + (1-\lambda)g_i(\vec{x}_2) \leq 0$
Тогда $\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2 \in X$ по определению X.
X — пересечение выпуклых множеств, а значит, выпукло ■.

Функция Лагранжа в ЗВП

$$f(\vec{x}) + \sum_{i=1}^{m} \lambda_i g(\vec{x}) = f(\vec{x}) + \big(\lambda, g(\vec{x})\big); \lambda_\iota \geq 0$$

Теорема Каруша-Джона:

$$\nabla f(\vec{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\vec{x}^*) = 0; \lambda_i g_i(\vec{x}^*) = 0; i = 1, \dots, m$$

Если X — выпукло, условие линейной независимости $\nabla g_i(\vec{x})$ можно заменить на **условия регулярности**

a) $\forall i \in 1, \dots, m$: $\exists \vec{x}_i \in X$, т.ч. $g_i(\vec{x}_i) < 0$ - **условие регулярности**
b) $\exists \vec{x} \in X$, т.ч. $\forall i \in 1 \dots m$: $g_i(\vec{x}) < 0$ - **условие регулярности Слейтера**

Очевидно, что $b \Rightarrow a$.
Пусть верно a. Поскольку X выпукло, можно выбрать $\vec{x}$:

$$\vec{x} := \sum_{i=1}^{m} \lambda_i \vec{x}_j, \sum_{i=1}^{m} \lambda_i = 1; \lambda_i \geq 0, i = 1 \dots m$$

В таком случае из b по неравенству Йенсена следует a.
Получается $a \Leftrightarrow b$.

Седловая точка
$(\vec{x}^*, \vec{y}^*)$ — седловая точка функции ψ аргументов $(\vec{x}, \vec{y})$ на $X \times Y$, если
$\psi(\vec{x}^*, y) \leq \psi(\vec{x}^*, \vec{y}^*) \leq \psi(\vec{x}, \vec{y}^*)$ при $\forall \vec{x} \in X, \vec{y} \in Y$

Теорема о седловой точке
Пусть функция Лагранжа ЗВП имеет седловую точку, т.е.
$\forall x \in \mathbb{R}^n, \lambda_i \geq 0, i = 1 \dots m$.

$$f(\vec{x}^*) + \sum_{i=1}^{m} \lambda_i g_i(\vec{x}^*) \leq f(\vec{x}^*) + \sum_{i=1}^{m} \lambda_i^* g_i(\vec{x}^*) \leq f(\vec{x}) + \sum_{i=1}^{m} \lambda_i^* g_i(\vec{x})$$

Тогда $\vec{x}^*$ — оптимальная точка ЗВП.

Доказательство
По левому неравенству:

$$\sum_{i=1}^{m} \lambda_i g_i(\vec{x}^*) \leq \sum_{i=1}^{m} \lambda_i^* g_i(\vec{x}^*)$$

По определению X: $\lambda_i^* \geq 0, g_i(\vec{x}^*) \leq 0$.
Так как λ — любое, при $\lambda = 0$:

$$0 \leq \sum_{i=1}^{m} \lambda_i^* g_i(\vec{x}^*) \Rightarrow \left(\lambda^*, g(\vec{x}^*)\right) = 0$$

Из правого неравенства $\forall \vec{x} \in X$:

$$f(\vec{x}^*) + 0 \leq f(\vec{x}) + \sum_{i=1}^{m} \lambda_i^* g_i(\vec{x}^*) \leq f(\vec{x})$$

По определению оптимальной точки $\vec{x}^*$ — оптимальна.

Теорема Куна-Таккера.
Пусть в ЗВП выполнено условие регулярности Слейтера. Тогда для того, чтобы x^* была оптимальной точкой ЗВП, необходимо и достаточно, чтобы для некоторого вектора λ^* с неотрицательными компонентами точка (x^*, λ^*) была седловой точкой функции Лагранжа.

В частности, если ψ — дифференцируема, то

$$\begin{cases} \dfrac{\partial\psi^*}{\partial\vec{x}} = 0 \\ \dfrac{\partial\psi^*}{\partial\lambda} \le 0 \\ \left(\lambda^*, \dfrac{\partial\psi^*}{\partial\vec{\lambda}}\right) = 0 \\ \lambda^* \ge 0 \end{cases}, \text{где } \begin{matrix} \dfrac{\partial\psi^*}{\partial\vec{x}} \stackrel{\text{def}}{=} \left.\dfrac{\partial\psi(\vec{x},\lambda)}{\partial\vec{x}}\right| \vec{x} = \vec{x}^*, \vec{\lambda} = \vec{\lambda}^* \\ \dfrac{\partial\psi^*}{\partial\vec{\lambda}} \stackrel{\text{def}}{=} \left.\dfrac{\partial\psi(\vec{x},\lambda)}{\partial\vec{\lambda}}\right| \vec{x} = \vec{x}^*, \vec{\lambda} = \vec{\lambda}^* \end{matrix}$$

В частности, если на $\vec{x}$ наложены координатные органичения ($\vec{x} \ge 0$), то данные условия приобретают вид:

$$\begin{cases} \dfrac{\partial\psi^*}{\partial\vec{x}} \ge 0; \left(\vec{x}^*, \dfrac{\partial\psi^*}{\partial\vec{x}}\right) = 0; \vec{x}^* \ge 0 \\ \dfrac{\partial\psi^*}{\partial\lambda} \le 0; \left(\lambda^*, \dfrac{\partial\psi^*}{\partial\lambda}\right) = 0; \lambda \ge 0 \end{cases}$$

Доказательство

Достаточность - из теоремы о седловой точке

Необходимость - без доказательства

Методы условной минимизации

Метод проекции градиента

Обобщение градиентного метода для задачи условной минимизации с выпуклым допустимым множеством. Так как возможен выход за пределы допустимого множества, то вводится операция проектирования на X (поиск ближайшей точки на X).

$\vec{x}^{(k+1)} = p_x\left(\vec{x}^{(k)} - \gamma\nabla f\left(\vec{x}^{(k)}\right)\right)$, где p_x – проектор на X

Метод условного градиента

В очередной точке $\vec{x}^{(k)}$ линеаризуют функцию $f(\vec{x})$.

Затем решают задачу минимизации линейной функции на X и найденную точку $\vec{x}^{(k)}$ используют для выбора направления движения.

$$\begin{cases} \overline{\vec{x}^{(k)}} = \underset{X}{argmin}\left(\nabla f\left(\vec{x}^{(k)}\right), \vec{x}\right) \\ \vec{x}^{(k+1)} + \gamma_k\left(\overline{\vec{x}^{(k)}} - \vec{x}^{(k)}\right) \end{cases}$$

Предполагается, что:

1. Задача минимизации линейной функции на X имеет решение
2. Это решение может быть найдено достаточно просто, лучше всего в явной форме
3. Нужно указать правило выбора γ_k. Значение γ_k можно определить из условия наискорейшего спуска:
   $\gamma = \underset{0\le\gamma\le1}{argmin}\, f\left(\vec{x}^{(k)} + \gamma\left(\vec{x}^{(k)} - \overline{\vec{x}^{(k)}}\right)\right)$

В этом случае последовательность $\vec{x}^{(k)}$ сходится к стационарной точке. В частности, для гладких функций f верно: $f(\vec{x}^*) - f^* = o\left(\frac{1}{k}\right)$, где $f^* = \min f(x)$ на множестве X.

Метод модифицированной функции Лагранжа

Функция Лагранжа в ЗВП:

$$\psi\left(\vec{x}, \vec{\lambda}\right) = f(\vec{x}) + \sum_{i=1}^{m} \lambda_i g_i(\vec{x}) = f(\vec{x}) + \left(\lambda, g(\vec{x})\right); \lambda_i \geq 0$$

По теореме о седловой точке, если:

$$\forall \vec{x} \in \mathbb{R}^n, \lambda \geq 0: \psi\left(\vec{x}^*, \vec{\lambda}\right) \leq \psi\left(\vec{x}^*, \vec{\lambda}^*\right) \leq \psi\left(\vec{x}, \vec{\lambda}^*\right)$$

То x^* — оптимальная точка задачи выпуклого программирования.

Это можно записать иначе:

$$\psi(x^*, \lambda^*) = \min_{x \in \mathbb{R}^n} \max_{\lambda \geq 0} \psi(\vec{x}, \vec{\lambda}) = \max_{\lambda \geq 0} \min_{x \in \mathbb{R}^n} \psi(\vec{x}, \vec{\lambda}) = f(\vec{x}^*)$$

Если $\vec{x}$ назвать **прямыми** переменными, а λ — двойственными, то видно, что прямые и двойственные переменные равноправны.

По теореме Куна-Таккера, исходную задачу можно заменить задачей отыскания седловой точки функции Лагранжа, т.е. задачи вида

$$\max_{\lambda \geq 0} \min_{x \in \mathbb{R}^n} \psi(\vec{x}, \vec{\lambda})$$

Модифицированная функция Лагранжа -

$$\mu\left(\vec{x}, \vec{\lambda}, k\right) = f(\vec{x}) + \frac{1}{2k} \left\| \vec{\lambda} + k g(\vec{x})_+ \right\|^2 - \frac{\|\lambda\|^2}{2k}$$

Здесь k - некоторый параметр (штраф); + - взятие положительной часть.

Итерационная формула для вычисления $\{\vec{x}^{(k)}, \vec{\lambda}^{(k)}\}$ имеет вид:

$$\begin{cases} \vec{x}^{(k+1)} = \arg\min \mu\left(\vec{x}, \vec{\lambda}^{(k)}, k\right), \vec{x} \in \mathbb{R}^n \\ \lambda^{(k+1)} = \left[\vec{\lambda}_k + \gamma_k \nabla \lambda \mu\left(\vec{x}^{(k+1)}, \vec{\lambda}^{(k)}, k\right)\right]_+ \end{cases}$$

Метод сходится к $(\vec{x}^*, \vec{\lambda}^*)$ со скорость геометрической процессии.

Метод Эрроу-Гурвица

$$\begin{cases} \vec{x}^{(k+1)} = \vec{x}^{(k)} - \gamma \frac{\partial \psi}{\partial \vec{x}}\left(\vec{x}^{(k)}, \vec{\lambda}^{(k)}\right) = \vec{x}^{(k)} - \gamma\left(\nabla f\left(\vec{x}^{(k)}\right) + \left(\nabla g\left(\vec{x}^{(k)}\right), \lambda^{(k)}\right)\right) \\ \lambda^{(k+1)} = \lambda^{(k)} + \gamma \frac{\partial \psi}{\partial \lambda}\left(\vec{x}^{(k)}, \vec{\lambda}^{(k)}\right) = \vec{\lambda}^{(k)} + \gamma g\left(\vec{x}^{(k)}\right) \end{cases}$$

Метод штрафных функций

Идея метода - сведение задачи с ограничениями к последовательности задач без ограничений

$$\min f(\vec{x}), \vec{x} \in X, X = \{\vec{x} \in \mathbb{R}^n, g_i(\vec{x}) \leq 0, i = 1 \dots m\}$$

Функция $\psi(\vec{x})$, определённая и непрерывная всюду в $\mathbb{R}^n$, называется штрафной функцией для рассматриваемой задачи с ограничениями, если

$$\begin{cases} \psi(\vec{x}) = 0, \vec{x} \in X \\ \psi(\vec{x}) > 0, \vec{x} \in \dfrac{\mathbb{R}^n}{X} \end{cases}$$

Строится однопараметрическое семейство функций:

$\psi(\vec{x}, \beta) = f(\vec{x}) + \frac{1}{\beta}\psi(\vec{x})$, где $\beta -$ скалярный параметр, принимающий строго положительные значения.

Алгоритм

Выберем такую убывающую последовательность $\{\beta_k\}_{k=1}^{\infty}$ положительных чисел, что $\lim_{k\to\infty} \beta_k = 0$. Сопоставим каждому β_k соответствующую функцию семейства $\psi(\vec{x}, \beta)$. Получаем последовательность функций $\psi(\vec{x}, \beta_1), \dots, \psi(\vec{x}, \beta_k)$.

Пусть для каждой функции этой последовательности может быть решена задача условной минимизации:

$\arg\min \psi(\vec{x}, \beta) = \vec{x}_\beta^*, \vec{x} \in \mathbb{R}^n$

Оказывается, что при некоторых условиях последовательность оптимальных точек для задач без ограничений к оптимальной точке для исходной задачи с ограничениями:

$\vec{x}_\beta^* \xrightarrow[\beta\to 0]{} \vec{x}^*$

$x^* = \arg\min f(\vec{x})$

Наиболее распространенная следующая штрафная функция:

$$\psi(\vec{x}) = \sum_{i=1}^{m} \left(g_i^+(\vec{x})\right)^2$$

Здесь $g_i^+(\vec{x}) = \max\{0, g_i(\vec{x})\}$

<u>Двойственность задачи выпуклого программирования</u>

В теореме Куна-Таккера прямые и двойственные переменные используются симметричным образом. Поэтому можно ожидать, что аналогичная симметрия существует и для задач оптимизации относительно прямых и двойственных переменных.

Введем функцию: $g(\vec{x}) = \sup \psi(\vec{x}, \lambda)\,, \lambda \geq 0.$

Тогда очевидно, что $g_i(\vec{x}) = f(\vec{x}) \leq 0; i = 1, \dots, m$

$g(\vec{x}) = \infty -$ в противном случае.

Понятно, что $\psi\left(\vec{x}, \vec{\lambda}\right) = f(\vec{x}) + \left(\lambda, g(\vec{x})\right); \lambda \geq 0$

В таком случае исходная ЗВП может быть записана в виде:

$\min g(\vec{x}) -?\ \forall \vec{x} \in \mathbb{R}^n$. Эту задачу принято называть **прямой.**

Поступим аналогичным образом, поменяв роль переменных и операций

max и min. Обозначим $h(\lambda) = \inf \Psi(\vec{x}, \vec{\lambda})$ при $\vec{x} \in \mathbb{R}^n$. Задачу нахождения максимума функции $h\left(\vec{\lambda}\right)-?$ при $\lambda \geq 0$ называю **двойственной.**

Теорема двойственности

Справедливы следующие соотношения двойственности

1) $f(\vec{x}) \geq h\left(\vec{\lambda}\right); \forall \vec{x} \in X, \lambda \geq 0$
2) Если выполнено условие теоремы Куна-Таккера, а пара (x^*, λ^*) есть седловая точка ф-и Лагранжда, то λ^* — решение двойственной задачи: $\lambda^* = \arg\max h(\vec{\lambda})$ при $\vec{\lambda} \geq 0$ и $f(\vec{x}^*) = h\left(\vec{\lambda}^*\right)$
3) Если для допустимых $\vec{x}^*, \vec{\lambda}^*$: $f(\vec{x}^*) = h\left(\vec{\lambda}^*\right)$, то
   - $\vec{x}^* = \arg\min f(\vec{x})$ при $\vec{x} \in X$ — решение прямой задачи
   - $\vec{\lambda}^* = \arg\max h(\vec{\lambda})$ при $\vec{\lambda} \geq 0$ — решение двойственной задачи

Доказательство

1) $f(\vec{x}) \geq f(\vec{x}) + \left(\vec{\lambda}, g(\vec{x})\right) = \Psi\left(\vec{x}, \vec{\lambda}\right) \geq \inf \Psi\left(\vec{x}, \vec{\lambda}\right) = h(\vec{\lambda})$
2) $\forall \lambda \geq 0$:
   $h(\lambda^*) = inf\, \Psi(\vec{x}, \vec{\lambda}^*) = \Psi\left(\vec{\mathrm{x}}^*, \vec{\lambda}^*\right) \geq \Psi\left(\vec{\mathrm{x}}^*, \vec{\lambda}\right) \geq \inf \Psi\left(\vec{\mathrm{x}}, \vec{\lambda}\right) = h(\vec{\lambda})$
   Отсюда $\vec{\lambda}^*$ — решение двойственной задачи. Но $\Psi\left(\vec{x}^*, \vec{\lambda}^*\right) = f(\vec{x}^*) \Rightarrow$ $h(\lambda^*) = f(x^*)$
3) На основании п.1 $f(\vec{x}) \geq h(\lambda^*) \underset{\text{по п.2}}{=} f(\vec{x}^*) \geq h(\vec{\lambda})$
   Тогда $\vec{x}^*$ — прямое решение, $\vec{\lambda}^*$ — двоственное решение.

Двойственность задачи линейного программирования

Рассмотрим множество $X = \{\vec{x} \in \mathbb{R}, \vec{x} \geq 0, A\vec{x} \geq b\}$.
$\vec{b} = (b_1, b_2, \ldots, b_m)^T, A$ — матрица $m \times n$; $f(\vec{x}) = (\vec{c}, \vec{x})$ — целевая функция (линейная).
ЗЛП: $\min f(\vec{x})-?$ при $\vec{x} \in \mathrm{X}$ — прямая задача линейного программирования
Построим функцию Лагранжа:
$\Psi\left(\vec{x}, \vec{\lambda}\right) = (\vec{c}, \vec{x}) + \left(\vec{\lambda}_1, \vec{b} - A\vec{x}\right) + \left(\vec{\lambda}_2 - \vec{x}\right), \vec{\lambda}_1 \in \mathbb{R}^m, \lambda_2 \in \mathbb{R}^n$
(удовлетворяет требованию $g_i(\vec{x}) \leq 0$).
Тогда $\min(\vec{c}, \vec{x}) = \max\inf((\vec{c}, \vec{x}) + \left(\vec{\lambda}_1, \vec{b} - A\vec{x}\right) + (\lambda_2, -\vec{x}))$ при $\vec{\lambda}_1 \geq$ $0, \vec{\lambda}_2 \geq 0, x \in \mathbb{R}^n = max\, inf\left(\left(\vec{c} - A^T\vec{\lambda}_1 - \lambda_2, \vec{x}\right) + \left(\vec{b}, \vec{\lambda}_1\right)\right) =$

$$= \max\begin{cases} -\infty, \text{если } \vec{c} - A^T\lambda_1 - \lambda_2 \neq 0 \\ (b, \lambda_1) \text{ если } \vec{c} = \vec{A}^T\vec{\lambda}_1 + \lambda_2 \end{cases} \text{при } \vec{\lambda}_1 \geq 0, \vec{\lambda} \geq 0$$