§2. Конечномерные гладкие задачи с равенствами

2.1 Постановка задачи

Пусть $f_i\colon \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \ldots, m$. Считаем, что все функции f_i обладают определенной гладкостью. Гладкая конечномерная экстремальная задача с ограничениями типа равенств:

$$f_0(x) \to \text{extr}; \quad f_i(x) = 0, \; i = 1, \ldots, m. \qquad (P)$$

Допустимые точки — удовлетворяющие ограничениям задачи. Функция Лагранжа задачи (P): $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$, вектор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$.

2.2.2 Конечномерная теорема об обратной функции

Теорема (конечномерная теорема об обратной функции)

Пусть $F\colon \mathbb{R}^n \to \mathbb{R}^n$, $F \in C^1(U)$, где $U \subset \mathcal{O}(\hat{x}, \mathbb{R}^n)$, $F(\hat{x}) = \hat{y}$ и якобиан отображения F в точке $\hat{x}$ отличен от нуля $\left(\det F'(\hat{x}) = \det\left(\frac{\partial F_i(\hat{x})}{\partial x_j}\right)_{i,j=1}^{n} \neq 0\right)$. Тогда $\exists$ обратное отображение F^{-1} некоторой окрестности V точки $\hat{y}$ в окрестность точки $\hat{x}$ такое, что $F^{-1}(\hat{y}) = \hat{x}$ и

$$F(F^{-1}(y)) = y, \quad |F^{-1}(y) - F^{-1}(\hat{y})| \leq K|y - \hat{y}| \quad \forall\, y \in V$$

с некоторой константой $K > 0$.

2.2. Необходимые и достаточные условия экстремума
2.2.1 Принцип Лагранжа

$$f_0(x) \to \mathrm{extr}; \quad f_i(x) = 0,\ i = 1, \dots, m. \qquad (P)$$

Функция Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$.

Теорема

Пусть $\hat{x} \in \mathrm{locextr}\, P$, $f_i \in C^1(U_{\hat{x}})$, $i = 0, 1, \dots, m$. Тогда $\exists\ \lambda \neq 0$: для $\mathscr{L}$ выполняется условие стационарности

$$\mathscr{L}'(\hat{x}) = 0 \quad \left(\Leftrightarrow \frac{\partial \mathscr{L}(\hat{x})}{\partial x_j} = 0,\ j = 1, \dots, n \Leftrightarrow \sum_{i=0}^{m} \lambda_i f_i'(\hat{x}) = 0 \right).$$

Точки, удовлетворяющие условию стационарности, называются стационарными.

Теорема

$\hat{x} \in \operatorname{locextr} P,\ f_i \in C^1(U_{\hat{x}}),\ i = 0, ..., m \Rightarrow \exists\ \lambda \neq 0 : \sum\limits_{i=0}^{m} \lambda_i f_i'(\hat{x}) = 0.$

$\lhd$ Доказательство проведем от противного. Предположим, что условие стационарности не выполняется. Это означает, что векторы $f_i'(\hat{x}) = \left(\dfrac{\partial f_i(\hat{x})}{\partial x_1}, \dots, \dfrac{\partial f_i(\hat{x})}{\partial x_n}\right)$, $i = 0, 1, \dots, m$, линейно независимы. Поэтому ранг матрицы

$$A = \left(\frac{\partial f_i(\hat{x})}{\partial x_j}\right)_{\substack{i=0,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} \dfrac{\partial f_0(\hat{x})}{\partial x_1} & \dots & \dfrac{\partial f_0(\hat{x})}{\partial x_n} \\ \dots & \dots & \dots \\ \dfrac{\partial f_m(\hat{x})}{\partial x_1} & \dots & \dfrac{\partial f_m(\hat{x})}{\partial x_n} \end{pmatrix}$$

равен $m+1$. Тогда по теореме о ранге матрицы существует матрица M порядка $(m+1) \times (m+1)$ с определителем, отличным от нуля.

Теорема

$\hat{x} \in \mathrm{locextr}\, P$, $f_i \in C^1(U_{\hat{x}})$, $i = 0, ..., m \Rightarrow \exists\ \lambda \neq 0 : \sum_{i=0}^{m} \lambda_i f_i'(\hat{x}) = 0$.

Допустим для определенности, что этой матрицей является матрица, составленная из первых столбцов матрицы A:

$$\det \begin{pmatrix} \dfrac{\partial f_0(\hat{x})}{\partial x_1} & \cdots & \dfrac{\partial f_0(\hat{x})}{\partial x_{m+1}} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial f_m(\hat{x})}{\partial x_1} & \cdots & \dfrac{\partial f_m(\hat{x})}{\partial x_{m+1}} \end{pmatrix} = \det M \neq 0.$$

Не ограничивая общности, считаем, что $f_0(\hat{x}) = 0$. Действительно, если $f_0(\hat{x}) \neq 0$, то следует рассмотреть функцию $\widetilde{f_0}(x) = f_0(x) - f_0(\hat{x})$ и для нее будет выполняться условие $\widetilde{f_0}(\hat{x}) = f_0(\hat{x}) - f_0(\hat{x}) = 0$, а условия стационарности и точки экстремума для функций f_0 и $\widetilde{f_0}$ одинаковы.

Для вектора $\bar{x} = (x_1, \ldots, x_{m+1})$ положим
$F(\bar{x}) = \big(F_0(\bar{x}), \ldots, F_m(\bar{x})\big) = \big(f_0(\bar{x}, \hat{x}_{m+2}, \ldots, \hat{x}_n), \ldots, f_m(\bar{x}, \hat{x}_{m+2}, \ldots, \hat{x}_n)\big)$.
Тогда функция F отображает некоторую окрестность точки $\hat{\bar{x}} = (\hat{x}_1, \ldots, \hat{x}_{m+1}) \in \mathbb{R}^{m+1}$ в $\mathbb{R}^{m+1}$, является (в силу условий гладкости теоремы) непрерывно дифференцируемым отображением в этой окрестности и $F(\hat{\bar{x}}) = 0$ в силу допустимости точки $\hat{x}$. Кроме того, якобиан отображения F в точке $\hat{\bar{x}}$ отличен от нуля, т. е.

$$\det \left(\frac{\partial F_i(\hat{\bar{x}})}{\partial x_j} \right)_{\substack{i=0,1,\ldots,m \\ j=1,\ldots,m+1}} = \det \left(\frac{\partial f_i(\hat{x})}{\partial x_j} \right)_{\substack{i=0,1,\ldots,m \\ j=1,\ldots,m+1}} = \det M \neq 0 \ .$$

По т. об обратной функции $\exists\ F^{-1}\colon U_0 \to U_{\hat{\bar{x}}}\ :\ F^{-1}(0) = \hat{\bar{x}}$ и $|F^{-1}(y) - F^{-1}(0)| \le K|y - 0| \Leftrightarrow |F^{-1}(y) - \hat{\bar{x}}| \le K|y|\ (K > 0)$.
В частности, для достаточно малого по модулю ε определен $\bar{x}(\varepsilon) := F^{-1}(\varepsilon, 0, \dots, 0)$, для которого $|\bar{x}(\varepsilon) - \hat{\bar{x}}| \le K|\varepsilon|$.
Тогда $F(\bar{x}(\varepsilon)) = (\varepsilon, 0, \dots, 0) \Leftrightarrow f_0(\bar{x}(\varepsilon), \hat{x}_{m+2}, \dots, \hat{x}_n) = \varepsilon$, $f_i(\bar{x}(\varepsilon), \hat{x}_{m+2}, \dots, \hat{x}_n) = 0$, $i = 1, \dots, m$. Таким образом, для $x(\varepsilon) = (x_1(\varepsilon), \dots, x_{m+1}(\varepsilon), \hat{x}_{m+2}, \dots, \hat{x}_n)$

$$f_0(x(\varepsilon)) = \varepsilon, \quad f_i(x(\varepsilon)) = 0,\ i = 1, \dots, m,$$
$$|x(\varepsilon) - \hat{x}| = |\bar{x}(\varepsilon) - \hat{\bar{x}}| \le K|\varepsilon|.$$

Отсюда $\hat{x} \notin \operatorname{locextr} P$, ибо вблизи него $\exists\ x(\varepsilon) \in D(P)$, на которых $f_0(x(\varepsilon)) = \varepsilon \gtrless 0 = f_0(\hat{x})$, т. е. f_0 принимает значения как большие так и меньшие чем $f_0(\hat{x})$.
Получили противоречие с тем, что $\hat{x} \in \operatorname{locextr} P$.
Следовательно, наше предположение (противного) неверно и тем самым теорема доказана. $\triangleright$

Пример 1. $4x_1 + 3x_2 \to \text{extr}; \quad x_1^2 + x_2^2 = 1.$

Функция Лагранжа $\mathscr{L} = \lambda_0(4x_1 + 3x_2) + \lambda(x_1^2 + x_2^2 - 1)$.

$$\mathscr{L}_x = 0 \iff \begin{cases} \mathscr{L}_{x_1} = 0, \\ \mathscr{L}_{x_2} = 0, \end{cases} \iff \begin{cases} 4\lambda_0 + 2\lambda x_1 = 0, \\ 3\lambda_0 + 2\lambda x_2 = 0. \end{cases}$$

$\lambda_0 = 0 \Rightarrow \lambda \neq 0 \Rightarrow x_1 = x_2 = 0$ — не является допустимой.

$\lambda_0 \neq 0,\ \lambda_0 = -1 \Rightarrow x_1 = \frac{4}{2\lambda},\ x_2 = \frac{3}{2\lambda} \overset{x_1^2+x_2^2=1}{\Rightarrow}$

$$\frac{16}{4\lambda^2} + \frac{9}{4\lambda^2} = 1 \iff 25 = 4\lambda^2 \iff \lambda = \pm\frac{5}{2}.$$

Стационарные точки $\left(\frac{4}{5}, \frac{3}{5}\right), \left(-\frac{4}{5}, -\frac{3}{5}\right)$.

По т. Вейерштрасса absmax и absmin достигаются.

$f_0\left(\frac{4}{5}, \frac{3}{5}\right) = 5,\ f_0\left(-\frac{4}{5}, -\frac{3}{5}\right) = -5.$

Ответ. $\left(-\frac{4}{5}, -\frac{3}{5}\right) \in \text{absmin}$, $S_{\text{absmin}} = -5$; $\left(\frac{4}{5}, \frac{3}{5}\right) \in \text{absmax}$, $S_{\text{absmax}} = 5$.

Пример 2. $\langle Ax, x\rangle \to \text{extr}; \quad \langle x, x\rangle = 1$

$\big(x \in \mathbb{R}^n,\ A = (a_{ij})_{i,j=1}^{n}$ — симметрическая матрица$\big)$.
По т. Вейерштрасса решения задач на минимум и на максимум существуют, поскольку сфера
$S^{n-1} = \{x \in \mathbb{R}^n \mid |x|^2 = \langle x, x\rangle = 1\}$ компактна.
Функция Лагранжа $\mathscr{L} = \lambda_0\langle Ax, x\rangle + \lambda(\langle x, x\rangle - 1)$.
Условие стационарности $\mathscr{L}_x = 0 \Leftrightarrow \lambda_0 Ax + \lambda x = 0$
(на семинаре выписывали $\mathscr{L}_x[h] = 2\lambda_0\langle Ax, h\rangle + 2\lambda\langle Ax, h\rangle$)
$\lambda_0 = 0 \Rightarrow \lambda \neq 0 \Rightarrow x = 0$ — не допустима.
$\lambda_0 \neq 0, \lambda_0 = -1 \Rightarrow Ax = \lambda x$. Умножим на $x \Rightarrow \langle Ax, x\rangle = \lambda\langle x, x\rangle = \lambda$.
Таким образом, стационарные точки — собственные вектора матрицы A. Решение на минимум — собственный вектор, соответствующий наименьшему собственному значению, решение на максимум — собственный вектор, соответствующий наибольшему собственному значению.

2.5 Задача Аполлония

Древнегреческим ученым Аполлонием (III–II век до н. э.) в книге “Конические сечения” решается задача о проведении из некоторой точки к эллипсу самого длинного и самого короткого отрезка. При этом он решает даже более общую задачу, определяя все проходящие через эту точку отрезки, перпендикулярные к эллипсу. Решим задачу Аполлония с помощью принципа Лагранжа.

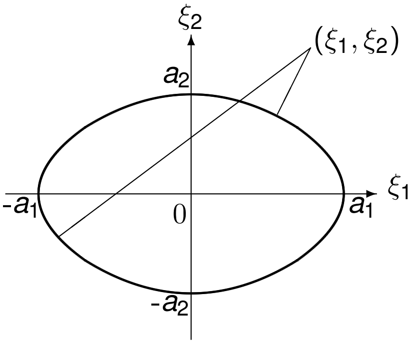

Рис. 1

Формализация:

$$f_0(x_1, x_2) = (x_1 - \xi_1)^2 + (x_2 - \xi_2)^2 \to \mathrm{extr};$$

$$f_1(x_1, x_2) = \left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 - 1 = 0 \quad (a_1 > a_2 > 0)$$

(здесь рассматривается эквивалентная задача об экстремуме квадрата расстояния).
Функция Лагранжа

$$\mathscr{L} = \lambda_0\left((x_1 - \xi_1)^2 + (x_2 - \xi_2)^2\right) + \lambda\left(\left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 - 1\right).$$

Условие стационарности:

$$\mathscr{L}_{x_i} = 0 \iff \lambda_0(x_i - \xi_i) + \lambda\frac{x_i}{a_i^2} = 0, \quad i = 1, 2.$$

$\lambda_0 = 0 \Rightarrow \lambda \neq 0 \Rightarrow x_1 = x_2 = 0 \notin$ эллипсу.
$\lambda_0 \neq 0,\ \lambda_0 = 1 \Rightarrow$

$$x_i-\xi_i+\frac{\lambda x_i}{a_i^2}=0 \Leftrightarrow x_i+\frac{\lambda x_i}{a_i^2}=\xi_i \Leftrightarrow \frac{x_i(a_i^2+\lambda)}{a_i^2}=\xi_i \Leftrightarrow x_i=\frac{\xi_i a_i^2}{a_i^2+\lambda}.$$

Подставим x_i в уравнение эллипса $\left(\frac{x_1}{a_1}\right)^2+\left(\frac{x_2}{a_2}\right)^2=1$:

$$\varphi(\lambda):=\frac{\xi_1^2 a_1^2}{(a_1^2+\lambda)^2}+\frac{\xi_2^2 a_2^2}{(a_2^2+\lambda)^2}=1. \qquad (1)$$

Относительно λ многочлен 4-ой степени $\Rightarrow$ число действительных корней ≤ 4 $\Rightarrow$ стационарных точек ≤ 4.

Каждому λ соответствует своя стационарная точка x.

По т. Вейерштрасса absmax и absmin достигаются.

Осталось решить (1), получить λ_j, найти $x(\lambda_j)$, $f_0\big(x(\lambda_j)\big)$ и найти наименьшее и наибольшее из полученных значений.

Если $\varphi(0)=\frac{\xi_1^2}{a_1^2}+\frac{\xi_2^2}{a_2^2}>1$, то (ξ_1,ξ_2) лежит вне эллипса.

Этот случай и будет изображен на рисунках.

Условия стационарности $x_i - \xi_i + \lambda \frac{x_i}{a_i^2} = 0 \Leftrightarrow \xi_i - x_i = \lambda \frac{x_i}{a_i^2}$,
геометрически означают, что вектор $(\xi - x) \parallel \operatorname{grad} f_1(x)$,
т. е. лежит на нормали к эллипсу.
Этот факт был впервые установлен Аполлонием.

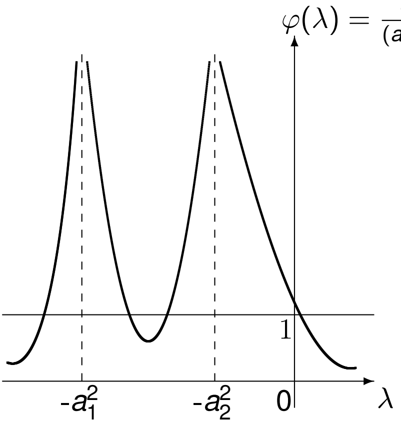

Рис. 2

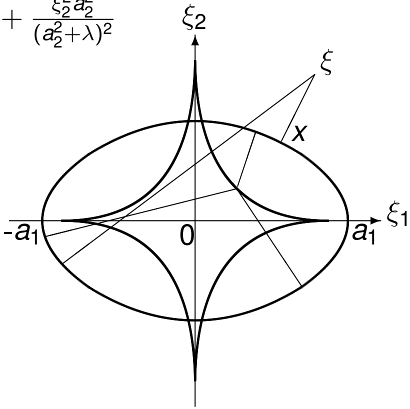

Рис. 3

$$\varphi(\lambda) := \frac{\xi_1^2 a_1^2}{(a_1^2+\lambda)^2} + \frac{\xi_2^2 a_2^2}{(a_2^2+\lambda)^2} = 1. \qquad (1)$$

Вывод уравнения "разделяющей" кривой (астроиды).

$$\varphi'(\lambda) = 0 \Leftrightarrow -\frac{2\xi_1^2 a_1^2}{(a_1^2+\lambda)^3} - \frac{2\xi_2^2 a_2^2}{(a_2^2+\lambda)^3} = 0 \Leftrightarrow (a_2^2+\lambda)^3(\xi_1 a_1)^2 =$$

$$= -(a_1^2+\lambda)^3(\xi_2 a_2)^2 \Leftrightarrow (a_2^2+\lambda)(\xi_1 a_1)^{2/3} = -(a_1^2+\lambda)(\xi_2 a_2)^{2/3}. \qquad (2)$$

Обозначим $a_1^2 + \lambda = A(\xi_1 a_1)^{2/3} \overset{(2)}{\Rightarrow} a_2^2 + \lambda = -A(\xi_2 a_2)^{2/3} \Rightarrow$
$a_1^2 - a_2^2 = A(\xi_1 a_1)^{2/3} + A(\xi_2 a_2)^{2/3} \Leftrightarrow A = \frac{a_1^2 - a_2^2}{(\xi_1 a_1)^{2/3} + (\xi_2 a_2)^{2/3}}$.
Подставим в (1) $a_1^2 + \lambda$, $a_2^2 + \lambda$, а затем и A:

$$\frac{\xi_1^2 a_1^2}{A^2(\xi_1 a_1)^{4/3}} + \frac{\xi_2^2 a_2^2}{A^2(\xi_2 a_2)^{4/3}} = 1 \Leftrightarrow (\xi_1 a_1)^{2/3} + (\xi_2 a_2)^{2/3} = A^2$$

$$\Leftrightarrow (\xi_1 a_1)^{2/3} + (\xi_2 a_2)^{2/3} = \frac{(a_1^2 - a_2^2)^2}{\left((\xi_1 a_1)^{2/3} + (\xi_2 a_2)^{2/3}\right)^2} \Leftrightarrow$$

$$\left((\xi_1 a_1)^{2/3} + (\xi_2 a_2)^{2/3}\right)^3 = (a_1^2 - a_2^2)^2 \Leftrightarrow (\xi_1 a_1)^{\frac{2}{3}} + (\xi_2 a_2)^{\frac{2}{3}} = (a_1^2 - a_2^2)^{\frac{2}{3}}.$$

Докажем, что касательная к астроиде $\perp$ к эллипсу.

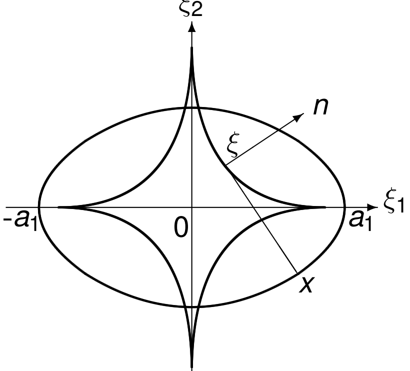

Рис. 4

Возьмем ξ — точку на астроиде. Пусть x — точка на эллипсе такая, что вектор $(x - \xi) \perp$ эллипсу. Для доказательства утверждения достаточно показать, что нормаль к астроиде — вектор $x - \xi$ является касательным, т. е. вектор $(x - \xi) \perp n$, где n — нормаль к астроиде в точке ξ.

Поскольку $x_i = \dfrac{\xi_i a_i^2}{a_i^2+\lambda}$, то

$$x-\xi = \left(\frac{\xi_1 a_1^2}{a_1^2+\lambda} - \xi_1, \frac{\xi_2 a_2^2}{a_2^2+\lambda} - \xi_2\right) = \left(-\frac{\lambda\xi_1}{a_1^2+\lambda}, -\frac{\lambda\xi_2}{a_2^2+\lambda}\right).$$

Нормалью к астроиде $(\xi_1 a_1)^{2/3} + (\xi_2 a_2)^{2/3} = (a_1^2 - a_2^2)^{2/3}$ является вектор пропорциональный градиенту функции $g(\xi_1,\xi_2) = (\xi_1 a_1)^{\frac{2}{3}} + (\xi_2 a_2)^{\frac{2}{3}}$, т. е. вектор $n = \left(\xi_1^{-\frac{1}{3}} a_1^{\frac{2}{3}}, \xi_2^{-\frac{1}{3}} a_2^{\frac{2}{3}}\right)$.

$$\langle x-\xi, n\rangle = -\frac{\lambda\xi_1\xi_1^{-\frac{1}{3}} a_1^{\frac{2}{3}}}{a_1^2+\lambda} - \frac{\lambda\xi_2\xi_2^{-\frac{1}{3}} a_2^{\frac{2}{3}}}{a_2^2+\lambda} = -\frac{\lambda(\xi_1 a_1)^{\frac{2}{3}}}{a_1^2+\lambda} - \frac{\lambda(\xi_2 a_2)^{\frac{2}{3}}}{a_2^2+\lambda} =$$

(подставим значения $a_i^2+\lambda$, выраженные через A и $(\xi_i a_i)^{2/3}$)

$$= -\frac{\lambda(\xi_1 a_1)^{2/3}}{A(\xi_1 a_1)^{2/3}} - \frac{\lambda(\xi_2 a_2)^{2/3}}{-A(\xi_2 a_2)^{2/3}} = -\frac{\lambda}{A} + \frac{\lambda}{A} = 0 \Rightarrow (x-\xi) \perp n.$$

$$f_0(x) \to \mathrm{extr}; \quad f_i(x) = 0,\ i = 1, \ldots, m \quad (f_i \colon \mathbb{R}^n \to \mathbb{R}). \qquad (P)$$

Теорема

$\hat{x} \in \operatorname{locmin} P$, $f_i \in D^2(\hat{x})$, $i = 0, 1, \ldots, m$ (условие гладкости), $\dim \operatorname{lin} \{f_1'(\hat{x}), \ldots, f_m'(\hat{x})\} = m$ (условие регулярности). Тогда $\exists\ \lambda = (1, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$: для $\mathscr{L}(x) = f_0(x) + \sum\limits_{i=1}^{m} \lambda_i f_i(x)$ выполняется условие стационарности

$$\mathscr{L}'(\hat{x}) = 0$$

и неотрицательной определенности матрицы вторых производных:

$$\langle \mathscr{L}''(\hat{x})h, h\rangle \geq 0 \quad \forall\ h \in K := \{\langle f_i'(\hat{x}), h\rangle = 0,\ i = 1, \ldots, m\}.$$

$$f_0(x) \to \mathrm{extr}; \quad f_i(x) = 0,\ i = 1, \ldots, m \quad (f_i \colon \mathbb{R}^n \to \mathbb{R}). \qquad (P)$$

Теорема

$f_i \in D^2(O(\hat{x}))$, $i = 0, 1, \ldots, m$ (условие гладкости),
$\dim \operatorname{lin} \{f_1'(\hat{x}), \ldots, f_m'(\hat{x})\} = m$ (условие регулярности),
$\exists\ \lambda = (1, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$: для $\mathscr{L}(x) = f_0(x) + \sum\limits_{i=1}^{m} \lambda_i f_i(x)$
выполняется условие стационарности

$$\mathscr{L}'(\hat{x}) = 0$$

и положительной определенности матрицы вторых производных:
$\langle \mathscr{L}''(\hat{x})h, h\rangle > 0\ \forall\ h \in K := \{\langle f_i'(\hat{x}), h\rangle = 0,\ i = 1, \ldots, m\},\ h \neq 0.$
Тогда $\hat{x} \in \operatorname{locmin} P$.