Теорема об обратной функции и примеры приложений теорем о неявной и обратной функциях.

Теорема (об обратной функции). *Если дифференциал гладкого отображения обратим во внутренний точке области определения отображения, то в некоторой окрестности этой точки обратимо и само отображение.*

Обратное отображение при этом определено в целой окрестности образа этой точки и при этом имеет ту же степень гладкости, что и исходное.

Дифференциал прямого и обратного отображений в соответствующих точках взаимно обратны.

Доказательство. Пусть G — область в пространстве X (например, в $\mathbb{R}^n$ или в $\mathbb{C}^n$), $f \in C^{(p)}(G, Y)$, $p \geq 1$, $x_0 \in G$, $y_0 = f(x_0)$ и $f'(x_0)$ — обратимо ($\|(f'(x_0))^{-1}\| \leq \infty$). (Заметим, что обратимость $f'(x_0)$ уже означает, что размерности пространств X и Y одинаковы).

Будем решать уравнение $y - f(x) = 0$ относительно x. Функция $F(x, y) = y - f(x)$ имеет ту же гладкость, что и f; $F(x_0, y_0) = y_0 - f(x_0) = 0$, а производная $F'_x(x, y) = -f'(x)$ непрерывна и обратима в (x_0, y_0).

По теореме о неявной функции найдутся окрестности $U(x_0)$, $V(y_0)$ и функция $x = \varphi(y)$ ($\varphi : V \to U$) такие, что $F(x, y) = y - f(x) = 0$ в $U \times V$ равносильно $x = \varphi(y)$ в V. Если окрестность $U(x_0)$ достаточно мала, то f можно считать определенным во всей окрестности U. Тогда получается, что функции (отображения) $f : U \to V$, $\varphi : V \to U$ взаимно обратные.

Из теоремы о неявной функции (см. там условия гладкости) и конкретного вида $F(x, y) = y - f(x)$ функции F вытекает, что функция $\varphi(y) = f^{-1}(y)$ имеет в V ту же степень гладкости, что и f в U.

Формула
$$\left(f^{-1}\right)'(y_0) = (f'(x_0))^{-1} \tag{1}$$
вытекает из правила дифференцирования неявной функции, если учесть, что в нашем случае $F'_y(x, y) \equiv E$.

Замечание. В случае, когда $X = Y = \mathbb{R}^n$ производная $f'(x)$ задается матрицей Якоби и ее обратимость в точке x_0 равносильна тому, что $\det f'(x_0) \neq 0$.

Отображения, непрерывные вместе с обратными к ним (которые предполагаются существующими) называют *гомеоморфизмами.* Отображения, гладкие (класса $C^{(p)}$, $p \geq 1$) вместе со своими обратными, называют *диффеоморфизмами* (данного класса гладкости).

Если нет специальных оговорок, то под диффеоморфизмом понимается диффеоморфизм класса $C^{(\infty)}$.

Замечание. Локальная обратимость отображения, вообще говоря не гарантирует его обратимости в целом.

Пример. (Полярные координаты). Как известно, полярные и декартовы координаты на плоскости (см. рис. 1) связаны соотношениями

$$\begin{cases} x = \rho\cos\varphi, \\ y = \rho\sin\varphi. \end{cases}$$

Якобиан перехода от полярных координат к декартовым равен ρ (проверьте), т.е. он положителен, если $\rho > 0$. Тем не менее, ввиду периодичности $\cos\varphi$ и $\sin\varphi$, это отображение не является обратимым в целом (даже в области, где $\rho > 0$), если рассматривать его как отображение из плоскости с координатами (ρ, φ) в плоскость с координатами (x, y).

Рис.1.

Задача.

1. Нарисуйте в плоскости (x, y) образы линий, изображенных на рис. 2.

2. В трехмерном пространстве полярные координаты называют *сферическими координатами.* Они могут быть введены в соответствии с рисунком 3. Запишите формулы, связывающие сферические координаты с декартовыми и найдите якобиан перехода.

Рис.2. Рис.3.

3. Полярные координаты в $\mathbb{R}^m$ вводятся соотношениями

$$\begin{cases} x^1 &= \rho\cos\varphi_1 \\ x^2 &= \rho\sin\varphi_1\cos\varphi_2 \\ \dots & \dots\dots\dots \\ x^{m-1} &= \rho\sin\varphi_1\sin\varphi_2\dots\sin\varphi_{m-2}\cos\varphi_{m-1} \\ x^m &= \rho\sin\varphi_1\sin\varphi_2\dots\sin\varphi_{m-2}\sin\varphi_{m-1} \end{cases}$$

покажите, что якобиан этого преобразования равен

$$\rho^{m-1}\sin^{m-2}\varphi_1\sin^{m-3}\varphi_2\dots\sin\varphi_{m-2}$$

и дайте геометрическую интерпретацию параметров $\rho,\varphi_1,\dots,\varphi_{m-1}$.

Если ρ,φ — полярные координаты на плоскости, то координатные линии $\rho = \mathrm{const}$, если их изображать в плоскости (x,y), вовсе не прямые, а кривые (окружности), поэтому полярные координаты являются примером, так называемых, криволинейных координат. Некоторые кривые (спирали, окружности) в этих координатах имеют более простую запись, чем в координатах (x,y). Сравните, $x^2+y^2=1$ и $\rho=1$.

Теорема об обратной функции позволяет для произвольных кривых (и вообще для кривых объектов) подбирать локальные криволинейные координаты в которых кривая (объект) записывается особенно просто — "выпрямляется".

Пример. ("Выпрямление" — см. рис. 4).

Рис.4.

Пусть на плоскости дана гладкая кривая $F(x,y)=0$. Пусть $\mathrm{rang}\, F' = \mathrm{rang}(F'_x, F'_y) = 1$ и пусть (x_0, y_0) ее точка, в которой, например, $F'_y(x_0, y_0) \neq 0$. Введем в окрестности (x_0, y_0) новые (криволинейные) координаты, связанные с данными (x,y) соотношениями

$$\begin{cases} \xi = x - x_0 \\ \eta = F(x,y) \end{cases}$$

По теореме об обратной функции, это отображение локально (в некоторой окрестности точки (x_0, y_0)) обратимо, т.к. его якобиан равен $F'_y(x,y)$ (проверьте). Значит, (ξ, η) могут служить координатами точек этой окрестности. На плоскости (x,y) координатные линии $\xi = \mathrm{const}$ и $\eta = \mathrm{const}$ — кривые (потому и название "криволинейные координаты"). Зато в плоскости (ξ, η) все наоборот и наша кривая $F(x,y)$ (точнее ее часть, попавшая в рассматриваемую окрестность) выпрямилась и задается самым простым способом $\eta = 0$.

Задача. (Поверхность размерности k в $\mathbb{R}^m$, различные формы ее определения и задания).

1. Если эталоном k-мерной поверхности в $\mathbb{R}^m$ считать подпространство заданное, например, соотношениями $x^{m-k+1} = \ldots = x^m = 0$ (свободными остались координаты $x^1, \ldots, x^k$, которые можно считать координатами $t^1, \ldots, t^k$ точек этого подпространства), то множество $S^k \subset \mathbb{R}^m$ естественно назвать *k-мерной поверхностью* (или *k-мерным многообразием*) в $\mathbb{R}^m$, если каждая точка S^k обладает окрестностью, которую можно диффеоморфно преобразовать (продеформировать) так (см. рис. 4), чтобы попавший в нее кусок S^k при этом перешел в кусок эталонной k-мерной поверхности (выпрямление).

Пусть $F(x) = (F^1, \ldots, F^{m-k})(x^1, \ldots, x^m)$ — набор гладких функций в $\mathbb{R}^m$, пусть $\mathrm{rang}\, F'(x) = m-k$. Покажите, что если система $F(x)=0$ разрешима, то она задает в $\mathbb{R}^m$ поверхность размерности k.

2. Закон движения точки в $\mathbb{R}^m$ имеет вид $x = x(t)$, где $x = (x^1, \dots, x^m)$, а параметр t — время. Если стоим на месте, то траектория просто одна точка, но если $\dot{x}(t) \neq 0$, т.е. если $\operatorname{rang} \dot{x}(t) = 1$, то траектория локально (в пределах небольшого интервала времени) — одномерная поверхность, то есть кривая в $\mathbb{R}^m$. Через достаточно большой промежуток времени мы можем вернуться в ту же точку пространства (траектория с самопересечениями), но локально (относительно t) траектория все-таки кривая.

Касательная к ней в точке $x_0 = x(t_0)$ имеет простую параметрическую (параметр τ) запись

$$x - x_0 = \dot{x}(t_0)\tau. \tag{2}$$

Если, например $\dot{x}^1(t_0) \neq 0$, то, выражая локально t через x^1 (обращая связь $x^1 = x^1(t)$), можем принять x^1 за независимый параметр. Подстановка $t = \varphi(x^1)$ в равенство $x = x(t)$ приведет к системе соотношений $x^1 = x^1$, $x^2 = x^2(\varphi(x^1)), \dots, x^m = x^m(\varphi(x^1))$ между $(x^1, \dots, x^m)$, из которых $m - 1$ нетривиальное. Эти соотношения локально задают нашу кривую в форме $F(x) = 0$, о которой говорилось в пункте 1. Задание кривой в виде $x = x(t)$ обычно называют *параметрическим заданием кривой.*

Пусть теперь $x = (x^1, \dots, x^m)$, $t = (t^1, \dots, t^k)$, $x(t)$ гладкая функция (отображение) и $\operatorname{rang} \dot{x}(t) = k$.

Покажите, что локально (относительно t) функция $x = x(t)$ осуществляет параметрическое задание k-мерной поверхности в $\mathbb{R}^m$ (и $t^1, \dots, t^k$ могут служить криволинейными координатами ее точек).

Как получить задание этой поверхности в виде $F(x) = 0$, где $F = (F^1, \dots, F^{m-k})$ и $\operatorname{rang} F'(x) = m - k$?

3. Как, наоборот, имея поверхность, заданную уравнением $F(x) = 0$ (где $\operatorname{rang} F'(x) = m - k$), приняв некоторые из переменных $(x^1, \dots, x^m)$ в качестве свободных параметров $(t^1, \dots, t^m)$, локально получить параметрическую форму задания этой поверхности?

4. Покажите, что если имеются два задания $x = x(t)$ и $F(x) = 0$ одной и той же k-мерной поверхности в $\mathbb{R}^m$ (т.е. $\operatorname{rang} \dot{x}(t) = k$ и $\operatorname{rang} F'(x) = m - k$), то касательная k-мерная плоскость к ней в точке $x_0 = x(t_0)$ задается соответственно в форме (2) и (3)

$$F'(x_0)(x - x_0) = 0 \tag{3}$$

Укажите развернутую координатную запись соотношений (2),(3).

5. Покажите, что плоскость $T_{x_0}S^k$, касательная к поверхности S^k в точке x_0, состоит из всех векторов скорости $\dot{x}(0)$ к траекториям движущихся по S^k точек, и только из них, если $x = x(t) \in S^k$, $t \in \mathbb{R}$ закон движения, а $x(0) = x_0$.

Природа теорем об обратной и о неявной функциях состоит в том, что дифференциал, являясь линейным отображением, локально хорошо приближает породившее его нелинейное отображение, поэтому теоремам линейной алгебры отвечают локальные теоремы анализа.

В частности, приведению линейно алгебраических объектов к каноническому виду (за счет разумного выбора систем координат) в анализе отвечает такое же, но локальное, приведение соответствующих объектов к каноническому (простейшему) виду. Один пример с "выпрямлением" объектов мы уже видели.

В качестве второго примера рассмотрим и докажем следующую полезную теорему о локальном приведении к каноническому виду гладкого отображения $f : \mathbb{R}^m \to \mathbb{R}^n$ в окрестности точки, где $\operatorname{rang} f'(x)$ постоянен.

Теорема (о ранге). *Пусть $f : U \to \mathbb{R}^n$ гладкое класса $C^{(p)}(U, \mathbb{R}^n)$, $p \geq 1$, отображение, определенное в окрестности $U \subset \mathbb{R}^m$ точки $x_0 \in \mathbb{R}^m$, а $y_0 = f(x_0)$. Если ранг отображения f (т.е. $\operatorname{rang} f'(x)$) в U постоянен и равен k, то существуют такие окрестности $O(x_0) \subset U$, $O(y_0) \subset \mathbb{R}^n$, точек x_0, y_0 и такие их диффеоморфизмы $u = \varphi(x)$, $v = \psi(y)$ класса $C^{(p)}$, что в окрестности $O(u_0) = \varphi(O(x_))$ точки $u_0 = \varphi(x_0)$, отображение $v = \psi \circ f \circ f^{-1}(u)$ (т.е. наше отображение f, записанное в координатах u и v, см. рис. 5) имеет вид*

$$(u^1, \ldots, u^k, \ldots, u^m) = u \mapsto v = (v^1, \ldots, v^n) = (u^1, \ldots, u^k, 0^1, \ldots, 0).$$

Рис.5.

Иными словами, теорема утверждает, что в окрестностях точек x_0 и $y_0 = f(x_0)$ вместо координат $(x^1, \ldots, x^m)$, $(y^1, \ldots, y^n)$ можно так ввести новые координаты $u^1, \ldots, u^m$, $v^1, \ldots, v^n$ соответственно, что в них отображение f будет иметь указанную простейшую запись.

Доказательство. Запишем координатное представление f и f'.

$$
\begin{array}{l}
y^1 = f^1(x^1,\dots,x^m),\\
\dots\dots\dots\\
y^n = f^n(x^1,\dots,x^m),
\end{array}
\quad
f'(x) = \begin{pmatrix}
\frac{\partial f^1}{\partial x^1} & \cdots & \cdots & \cdots & \frac{\partial f^1}{\partial x^m}\\
\vdots & \ddots & \vdots & & \vdots\\
\cdots & \cdots & \frac{\partial f^k}{\partial x^k} & \cdots & \cdots\\
\vdots & \ddots & & & \vdots\\
\frac{\partial f^n}{\partial x^1} & \cdots & \cdots & \cdots & \frac{\partial f^n}{\partial x^m}
\end{pmatrix}
$$

и без ограничения общности рассуждений (объясните!) положим, что ранг матрицы f' реализуется на указанном главном миноре порядка k.

Тогда рассмотрим отображение $u = \varphi(x)$, определяемое вместе с его производной $\varphi'(x)$ следующими соотношениями

$$
\begin{array}{l}
u^1 = f^1(x^1,\dots,x^m),\\
\dots\dots\dots\\
u^k = f^k(x^1,\dots,x^m),\\
u^{k+1} = x^{k+1},\\
u^m = x^m,
\end{array}
\quad
\varphi'(x) = \begin{pmatrix}
\frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^k} & \cdots & \cdots & \frac{\partial f^1}{\partial x^m}\\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots\\
\frac{\partial f^k}{\partial x^1} & \cdots & \frac{\partial f^k}{\partial x^k} & \cdots & \cdots & \frac{\partial f^k}{\partial x^m}\\
 & & & 1 & & \\
 & 0 & & & \ddots & \\
 & & & & & 1
\end{pmatrix}
$$

Поскольку $\det \varphi'(x_0) \neq 0$, по теореме об обратной функции отображение φ обратимо в некоторой окрестности $O(x_0)$ точки x_0. Отображение φ^{-1} при этом определено в соответствующей окрестности $O(u_0)$ точки $u_0 = \varphi(x_0)$ и имеет ту же степень гладкости, что и f.

В новых координатах $(u^1,\dots,u^m)$ отображение $y = f(x) = f\circ\varphi^{-1}(u) = g(u)$ и его производная $g'(u)$ имеют вид

$$
\begin{array}{l}
y^1 = u^1,\\
\dots\dots\dots\\
y^k = u^k,\\
y^{k+1} = g^{k+1}(u^1,\dots,u^m),\\
\dots\dots\dots\\
y^n = g^n(u^1,\dots,u^m),
\end{array}
\quad
g'(x) = \begin{pmatrix}
1 & & & & & \\
 & \ddots & & & 0 & \\
 & & 1 & & & \\
\frac{\partial g^{k+1}}{\partial u^1} & \cdots & \frac{\partial g^{k+1}}{\partial u^k} & \cdots & \cdots & \frac{\partial g^{k+1}}{\partial u^m}\\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots\\
\frac{\partial g^n}{\partial u^1} & \cdots & \frac{\partial g^n}{\partial u^k} & \cdots & \cdots & \frac{\partial g^n}{\partial u^m}
\end{pmatrix}
$$

Но $g'(u) = f'(x)\circ(\varphi^{-1})'(u)$, причем $\operatorname{rang} f'(x) = k$, а $\operatorname{rang}(\varphi^{-1})'(u) = m$, поэтому $\operatorname{rang} g'(u) = k$. Значит в матрице $g'(u)$ равны нулю все частные производные $\partial_i g^j$, $i = k+1,\dots,m$; $j = k+1,\dots,n$. Но это в свою очередь означает, что функции g^j, $j = k+1,\dots,n$ не зависят от переменных u^i, $i = k+1\dots,n$. Учитывая это ключевое наблюдение, запишем последнее отображение $v = \psi(y)$ и его производную

$$
\begin{array}{l}
v^1 = y^1,\\
\dots\dots\dots\\
v^k = y^k,\\
v^{k+1} = y^{k+1} - g^{k+1}(y^1,\dots,y^m),\\
\dots\dots\dots\\
v^n = y^n - g^n(y^1,\dots,y^m),
\end{array}
\quad
\psi'(x) = \begin{pmatrix}
1 & & & & & \\
 & \ddots & & & & 0 \\
 & & 1 & & & \\
-\frac{\partial g^{k+1}}{\partial y^1} & \cdots & -\frac{\partial g^{k+1}}{\partial y^k} & 1 & & \\
\vdots & \ddots & \vdots & & \ddots & \\
-\frac{\partial g^n}{\partial y^1} & \cdots & -\frac{\partial g^n}{\partial y^k} & & & 1
\end{pmatrix}
$$

Поскольку $\det \psi'(y) \equiv 1$, по теореме об обратной функции отображение ψ обратимо в некоторой окрестности $O(y_0)$ точки $y_0 = f(x_0)$. Отображение ψ^{-1}, определенное в соответствующей окрестности $O(v_0)$ точки $v_0 = \psi(y_0)$, имея ту же гладкость, что и g, т.е., что и f, вводит таким образом, в $O(y_0)$ новые координаты $(v^1,\dots,v^n)$.

В координатах $(u^1,\dots,u^m)$, $(v^1,\dots,v^n)$ для отображения $\psi\circ g = \psi\circ f\circ \varphi^{-1}$ теперь получаем указанное в формулировке теоремы представление.

Доказанная теорема, очевидно, тоже является локальным нелинейным вариантом соответствующей теоремы линейной алгебры, утверждающей, что любое линейное отображение $f : \mathbb{R}^m \to \mathbb{R}^n$ выбором новых базисов в $\mathbb{R}^m$ и $\mathbb{R}^n$, т.е. выбором должных систем координат, приводится к указанному виду.

Задача. (Канонические переменные и система уравнений Гамильтона)

1. В вариационном исчислении и фундаментальных вариационных принципах классической механики важную роль играет следующая **система уравнений Эйлера-Лагранжа:**

$$
\begin{cases}
\left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial v}\right)(t,x,v) = 0,\\
v = \dot{x}(t),
\end{cases}
$$

где $L(t,x,v)$ — заданная функция переменных t,x,v, среди которых t обычно является временем, x — координатой, а v — скоростью. Это система двух уравнений на три переменные. Из нее обычно желают найти зависимости $x = x(t)$ и $v = v(t)$, что по существу сводится к отысканию закона движения $x = x(t)$, ибо $v = \dot{x}(t)$.

Запишите подробно первые уравнения системы, раскрыв производную $\frac{d}{dt}$, с учетом того, что $x = x(t)$ и $v = v(t)$.

2. Покажите, что если от переменных t,x,v,L перейти к так называемым *каноническим* переменным t,x,p,H, сделав преобразование Лежандра

$$
\begin{cases}
p = \frac{\partial L}{\partial v},\\
H = pv - L
\end{cases}
$$

по переменным v, L, заменяя их на переменные p, H, то система Эйлера-Лагранжа приобретет симметричный вид

$$\dot{p} = -\frac{\partial H}{\partial x}, \qquad \dot{x} = \frac{\partial H}{\partial p}$$

3. В механике чаще всего вместо x и $v = \dot{x}$ используют обозначения q и $\dot{q}$. В многомерном случае, когда $L(t, q, \dot{q}) = L(t, q^1, \ldots, q^m, \dot{q}^1, \ldots, \dot{q}^m)$ система уравнений Эйлера-Лагранжа имеет вид

$$\left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\right)(t, q, \dot{q}) = 0 \quad (i = 1, \ldots, m)$$

Сделав преобразование Лежандра по переменным $\dot{q}, L$ перейдите от переменных $t, q, \dot{q}, L$ к каноническим переменным t, q, p, H и покажите, что при этом система уравнений Эйлера-Лагранжа перейдет в следующую систему уравнений Гамильтона

$$\dot{p}_i = -\frac{\partial H}{\partial q^i}, \qquad \dot{q}^i = \frac{\partial H}{\partial p_i} \ (i = 1, \ldots, m)$$

Задача. Покажите, что корни $z_1, \ldots, z_n$ уравнения $z^n + c_1 z^{n-1} + \cdots + c_n = 0$ гладко зависят от коэффициентов $c_1, \ldots, c_n$ (во всяком случае, пока все корни различны).

Задача. 1. "Если $f(x, y) = 0$, то $\frac{\partial y}{\partial x} \cdot \frac{\partial x}{\partial y} = 1$, а если $f(x, y, z) = 0$, то $\frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x} \cdot \frac{\partial x}{\partial z} = -1$." Придайте точный смысл этим высказываниям.

2. Проверьте второе из них на примере формулы Клапейрона $\frac{p \cdot v}{T} = \mathrm{const}$ и докажите его в общем случае функции трех переменных.

3. Как выглядит аналогичное высказывание для соотношения $f(x^1, \ldots, x^m) = 0$ между m переменными? Проверьте его справедливость.

Задача. (Преобразование Лежандра).

Элементы векторного пространства X и сопряженного к нему пространства X^* (состоящего из линейных на X функций) будем обозначать соответственно через x и x^*, а под спариванием (x^*, x) будем понимать значение $x^*(x)$ функции $x^* \in X^*$ на элементе $x \in X$.

Гладкой функции $f : X \to \mathbb{R}$ на X сопоставим функцию $f^* : X^* \to \mathbb{R}$ на X^* (называемую *преобразованием Лежандра функции* f) по правилу

$$f^*(x^*) \stackrel{\mathrm{def}}{=} \sup_x ((x^*, x) - f(x)) \qquad (*)$$

или в виде соотношений (замены переменных x, f на x^*, f^*)

$$\begin{cases} x^* = f'(x) \\ f^*(x^*) = (x^*, x) - f(x) \end{cases} \tag{$**$}$$

1. Какова взаимосвязь этих определений?

2. Найдите преобразование Лежандра функции $f : \mathbb{R} \to \mathbb{R}$, если $f(x) = x^p$ $(p > 1)$.

3. Проверьте выпуклость функции f^*.

4. Докажите неравенство $(x^*, x) \le f^*(x^*) + f(x)$ и запишите его в случае $f(x) = x^p$ $(p > 1)$.

5. Пусть f — выпуклая функция. Покажите, что определения $(*)$, $(**)$ могут быть переписаны в следующем симметричном относительно x, f, x^*, f^* виде

$$\begin{cases} x^* = f'(x), \qquad x = f^{*\prime}(x^*) \\ (x^*, x) = f^*(x^*) + f(x) \end{cases} \tag{$**'$}$$

Получите отсюда инволютивность преобразования Лежандра $((f^*)^*) = f$ и проверьте ее на примере функции $f(x) = x^p$ $(p > 1)$.

Задача. (Преобразование Лежандра функций многих переменных.)

Для функций $f : \mathbb{R}^m \to \mathbb{R}$ преобразование Лежандра имеет вид

$$\begin{cases} \xi_i = \frac{\partial f}{\partial x^i}(x^1, \ldots, x^m), \qquad i = 1 \ldots m \\ f^*(\xi^1, \ldots, \xi^m) = \sum_{i=1}^m \xi_i x^i - f(x^1, \ldots, x^m). \end{cases} \tag{$**''$}$$

1. Дайте геометрическую интерпретацию этого преобразования как перехода от координат $(x^1, \ldots, x^m, f(x^1, \ldots, x^m))$ точки на графике функции $f(x)$ к параметрам $(\xi^1, \ldots, \xi^m, f^*(\xi^1, \ldots, \xi^m))$, связанным с плоскостью, касательной к графику в этой точке.

2. Покажите, что преобразование $(**'')$ локально заведомо возможно, если $f \in C^{(2)}$ и $\det(\frac{\partial^2 f}{\partial x^i \partial x^j}) \ne 0$.

3. Проверьте, что

$$df^* = \sum_{i=1}^m x^i\, d\xi_i + \sum_{i=1}^m \xi_i\, dx^i - df = \sum_{i=1}^m x^i\, d\xi_i.$$

4. Мыльная пленка, натянутая на проволочный контур, образует так называемую *минимальную* поверхность, имеющую наименьшую (точнее, экстремальную) площадь среди всех поверхностей, натянутых на этот контур.

Если локально задать эту поверхность как график функции $z = f(x, y)$, то, оказывается, функция f должна удовлетворять следующему уравнению минимальных поверхностей

$$(1 + f_y')^2 f_{xx}'' - 2f_x' f_y' f_{xy}'' + (1 + f_x')^2 f_{yy}'' = 0$$

Покажите, что после преобразования Лежандра это уравнение приводится к линейному уравнению

$$(1 + \eta^2) {f^*}_{\eta\eta}'' + 2\xi\eta {f^*}_{\xi\eta}'' + (1 + \xi^2) {f^*}_{\xi\xi}'' = 0$$

Зависимость функций и другие примеры применения теоремы об обратной функции (лемма Морса, разложение диффеоморфизма).

Анализируя проведенное доказательство теоремы о ранге можно сделать два следующих полезных наблюдения.

Замечание 1. Если $k = n$, т.е. если ранг гладкого отображения $f : U \to \mathbb{R}^n$ равен размерности пространства образа, то точка $y_0 = f(x_0)$ является внутренней для $f(U)$.

Иными словами, образ $f(U)$ открытого множества U при таком отображении является открытым множеством в $\mathbb{R}^n$.

Доказательство. Действительно, если бы вместо x_0, y_0 и f речь шла об u_0, v_0 и $\psi \circ f \circ \varphi^{-1}$, то из явной записи отображения $\psi \circ f \circ \varphi^{-1}$ непосредственно видеть, что если параметры $(u^1, \ldots, u^k, \ldots, u^m)$ свободно меняются в окрестности $u_0 = (u_0^1, \ldots, u_0^m) = (0, \ldots, 0)$, то соответствующие параметры $(v^1, \ldots, v^n)$ при $k = n$ могут задать любую точку в некоторой окрестности $v_0 = (v_0^1, \ldots, v_0^n) = (0, \ldots, 0)$. Но при диффеоморфизмах φ и ψ внутренние точки переходят во внутренние (теорема об обратной функции), поэтому для отображения $f = \psi^{-1} \circ (\psi \circ f \circ \varphi^{-1}) \circ \varphi$ это свойство тоже имеет место.

Замечание 2. Если $k < n$, т.е. ранг гладкого отображения $f : U \to \mathbb{R}^n$ в окрестности точки x_0 постоянен и равен $k < n$, то найдется $n - k$ определенных в окрестности точки $y_0 = f(x_0)$ функций $g^i(y^1, \ldots, y^k)$, $i = k + 1, \ldots, n$, той же степени гладкости, что и f, зависящих только от k переменных, которые устанавливают следующие $n - k$ соотношений между задающими f функциями $f^1, \ldots, f^n$:

$$f^i(x^1, \ldots, x^m) = g^i(f^1(x^1, \ldots, x^m), \ldots, f^k(x^1, \ldots, x^m)) \quad (i = k + 1, \ldots, n)$$

Эти соотношения написаны в предположении (как и при доказательстве теоремы), что $\operatorname{rang} f'(x) = k$ реализуется на первых k функциях $f^1, \dots, f^k$. Таким образом, получается что подобно линейной зависимости векторов), остальные функции $f^{k+1}, \dots, f^n$ системы $f^1, \dots, f^n$ выражаются (правда, не обязательно линейно, но функционально) через функции $f^1, \dots, f^k$.

Доказательство. Написанные соотношения получаются, если в координатной записи отображения $y = g(u)$ учесть, что функции g^i, $i = k+1, \dots, n$ зависят только от $u^1, \dots, u^k$, при этом $u^1 = y^1, \dots, u^k = y^k$, а $y^j = f^j(x^1, \dots, x^m)$, $j = 1, \dots, n$.

Зависимость функций.

Естественно принять следующее определение.

Определение. *Система функций* $f^i(x)$ $(i = 1, \dots, n)$ функционально независима *в окрестности точки* x_0, *если для любой, определенной в окрестности точки* $y_0 = f(x_0)$, $y_0 = (y_0^1, \dots, y_0^n)$, *функции* $F(y) = F(y^1, \dots, y^n)$, *соотношение*

$$F(f^1(x), \dots, f^n(x)) \equiv 0$$

в окрестности x_0 *возможно только, если* $F(y) \equiv 0$ *в окрестности* y_0.

Линейная независимость, рассматриваемая в алгебре, есть независимость по отношению к линейным соотношениям

$$F(y^1, \dots, y^n) = \lambda_1 y^1 + \dots + \lambda_n y^n$$

Если система не является функционально независимой, то говорят, что она *функционально зависима.*

В случае линейной зависимости системы векторов некоторые из них оказываются линейными комбинациями других.

Аналогичная ситуация, как мы видели, имеет место и в случае Функциональной зависимости. Объединим сделанные выше замечания 1 и 2 и сформулируем полезное утверждение.

Утверждение. *Если система* $f^i(x^1, \dots, x^m)$ $(i = 1, \dots, n)$ *гладких функций, определенных в окрестности* $U(x_0)$ *точки* $x_0 \in \mathbb{R}^m$, *такова, что ранг матрицы*

$$\begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \dots & \frac{\partial f^1}{\partial x^m} \\ \vdots & & \vdots \\ \frac{\partial f^n}{\partial x^1} & \dots & \frac{\partial f^n}{\partial x^m} \end{pmatrix} (x)$$

в любой точке $x \in U$ один и тот же и равен k, то

a. *при $k = n$ система функционально независима в окрестности x_0.*

b. *при $k < n$ найдутся окрестность точки x_0 и такие k функций системы, пусть $f^1, \ldots, f^k$, что остальные $n - k$ функций системы представляются в виде*

$$f^i(x) = g^i(f^1(x), \ldots, f^k(x)),$$

где $g(y) = g(y^1, \ldots, y^k$ — гладкие функции, определенные в окрестности точки $y_0 = (f^(x_0), \ldots, f^n(y_0)$ и зависящие только от k координат текущей точки $y = (y^1, \ldots, y^n)$.

Упражнение. Опираясь на замечания 1 и 2, проведите все же подробное обоснование этого утверждения.

В качестве дальнейших примеров применения теоремы об обратной функции попробуйте теперь самостоятельно доказать два красивых и полезных утверждения, приведенных в следующих задачах.

Задача. (Локальное разложение диффеоморфизма в композицию простейших)

Диффеоморфизм $y = g(x)$ вида $y^i = x^i$ $(i \neq j)$, $y^j = g^j(x^1, \ldots, x^m)$, т.е. диффеоморфизм, при котором изменяется только одна координата, назовем *простейшим.* Докажите (например, по индукции), что любой диффеоморфизм локально представим в виде композиции простейших.

Задача. (Лемма Морса или приведение к каноническому виду функции в окрестности невырожденной критической точки).

Точка x_0 называется *критической точкой* гладкой функции $f : U(x_0) \to \mathbb{R}$, если $f'(x_0) = 0$. Критическая точка x_0 называется *невырожденной критической точкой*, если форма $f''(x_0)$ не вырождена, т.е. если $\det(\frac{\partial^2 f}{\partial x^i \partial x^j}) \neq 0$ (т.е. гессиан функции f в точке отличен от нуля).

По формуле Тейлора в окрестности критической точки x_0 имеет место разложение

$$f(x) - f(x_0) = \frac{1}{2}\frac{\partial^2 f}{\partial x^i \partial x^j}(x)(x^i - x_0^i)(x^j - x_0^j) + o(|x - x_0|^2)$$

Лемма Морса утверждает, что если x_0 невырожденная критическая точка, то найдется диффеоморфизм $x = \varphi(u)$, вводящий в окрестности точки $x_0 = (x_0^1, \ldots, x_0^m)$ новые координаты $(u^1, \ldots, u^m)$, в которых вся правая часть тейлоровского разложения приводится к следующему каноническому виду

$$f \circ \varphi(u) - f(x_0) = -\Big((u^1)^2 + \cdots + (u^k)^2\Big) + \Big((u^{k+1})^2 + \cdots + (u^m)^2\Big)$$

Указание. Вспомните процесс приведения квадратичной формы к каноническому виду и повторите его, применяя, где надо, теорему об обратной функции.

Условный экстремум

Одним из наиболее ярких и популярных достижений классического дифференциального исчисления являются предлагаемые им рецепты отыскания экстремумов функций. Необходимые и достаточные признаки экстремума, которые мы получили, относятся, как уже отмечалось, к внутренним экстремумам, когда аргумент x функции $f(x)$ может принимать любое значение в некоторой окрестности исследуемой точки.

Часто возникает более сложная и с практической точки зрения даже более интересная ситуация, когда имеется экстремум функции при некоторых условиях, ограничивающих область применения аргумента. Типичным примером может служить изопериметрическая задача, когда ищется тело, имеющее максимальный объем при условии, что ограничивающая тело поверхность имеет заданную площадь. Или, например, ищется наиболее экономная в том или ином смысле (по расходу топлива, по расстоянию, по времени движения и т.д.) траектория запуска пилотируемого корабля при условии, что перегрузки и ряд других показателей в полете не должны превышать определенных критических уровней. Такие задачи рассматривает теория оптимального управления. Вернемся, однако, к первому примеру в его наиболее простой и доступной нам сейчас форме.

Пусть среди всех прямоугольников с заданным периметром $2p$ ищется тот, у которого наибольшая площадь. Если x и y — величины сторон прямоугольника, то задача состоит в отыскании максимума функции $f(x,y) = xy$ при условии, что $x + y = p$. Эта задача, конечно, решается без труда, а ее обобщенный аналог, задача на условный экстремум, которую мы здесь рассмотрим, состоит в следующем.

Ищется экстремум вещественнозначной функции

$$f(x^1, \ldots, x^n) \tag{1}$$

при условии, что переменные связаны соотношениями

$$F_i(x^1, \ldots, x^n) = 0 \qquad i = 1, \ldots, m \tag{2}$$

Мы будем предполагать, что все рассматриваемые функции дифференцируемы и даже непрерывно дифференцируемы.

Если ранг системы функция (2) равен m, т.е. все условия в системе (2) существенны (независимы) и она совместна, то система задает гладкую $(n-m)$-мерную поверхность $S \subset \mathbb{R}^n$ и задача состоит в отыскании экстремумов функции $f|_S$ — ограничения (сужения) f на поверхность S.

Можно было бы теперь сформулировать задачу отыскания условного экстремума в следующем общем виде.

Дана вещественнозначная функция $f : X \to \mathbb{R}$ и подмножество $S \subset X$ области ее определения. Найти экстремумы функции $f|_S$.

1. Простейший случай.

Рассмотрим простейшую нетривиальную ситуацию, когда ищется точка (x_0, y_0) экстремума функции $f(x, y)$ двух переменных, связанных одним соотношением $F(x, y) = 0$. Считаем, что $\operatorname{rang} F' = 1$ и, конечно, $F(x_0, y_0) = 0$. Пусть, например, $F'_y(x_0, y_0) \neq 0$. Тогда по теореме о неявной функции все множество S, задаваемое условием $F(x, y) = 0$, в некоторой окрестности точки (x_0, y_0) можно считать графиком некоторой функции $y = \varphi(x)$. Значит, если (x_0, y_0) — точка экстремум для $f|_S$, то x_0 — точка внутреннего экстремума функции $f(x, \varphi(x))$. Тогда ее производная в точке x_0 должна быть нулевой. Запишем это необходимое условие, используя правило дифференцирования неявной функции

$$\frac{df}{dt} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{d\varphi}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\left(-\frac{\partial F}{\partial y}\right)^{-1}\left(\frac{\partial F}{\partial x}\right) = 0$$

полагая $\frac{\partial f}{\partial y}\left(\frac{\partial F}{\partial y}\right)^{-1} = \lambda$, приходим к следующему соотношению

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \lambda\left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}\right), \tag{3}$$

которое должно быть выполнено в точке (x, y), если она является точкой (x_0, y_0) экстремума нашей задачи.

Если к равенству (3) добавить уравнение связи $F(x, y) = 0$, то мы получим систему

$$\frac{\partial f}{\partial x} = \lambda\frac{\partial F}{\partial x},\ \frac{\partial f}{\partial y} = \lambda\frac{\partial F}{\partial y},\ F(x, y) = 0 \tag{4}$$

трех уравнений с тремя неизвестными, решая которую найдем точки, среди которых только и могут быть точки экстремума.

Если, следуя Лагранжу, написать вспомогательную функцию

$$L(x, y, \lambda) = f(x, y) - \lambda F(x, y), \tag{5}$$

то систему (4) можно получить как необходимое условие внутреннего экстремума функции L переменных x, y, λ.

Функцию L называют функцией Лагранжа, а описанный способ отыскания условного экстремума (т.е. составление функции Лагранжа (5) и запись необходимых условий экстремума в форме (4)) называется *методом Лагранжа* или, точнее, *методом множителей Лагранжа*. Почему вдруг множителей, а не множителя, ясно из следующего обобщения.

2. Обобщение.

Изложенное выше можно дословно повторить, считая, что $(x, y) = (x^1, \ldots, x^m, y^1, \ldots, y^k)$, что F — это вектор-столбец условий вида (2) (где вместо x переменная пока обозначена через (x, y)), что $\operatorname{rang} F' = k$ и что этот ранг реализуется, например, на производной F по переменным $y^1, \ldots, y^k$, т.е. F'_y — обратима.

В этом случае $\frac{\partial f}{\partial y}\left(\frac{\partial F}{\partial y}\right)^{-1} = \lambda$ окажется матрицей-строкой $(\lambda_1, \ldots, \lambda_k)$, а функция Лагранжа будет иметь вид

$$L = f - \lambda F = f - (\lambda_1 F_1 + \cdots \lambda_k F_k) \tag{5'}$$

Теперь можно забыть про какие-то выделенные по ходу рассуждений переменные y, на которых реализовывался ранг F' и, имея функцию (1) и условия (2), можно сразу составить симметричную относительно переменных функцию Лагранжа ($5'$) и после этого написать необходимые условия экстремума в нашей задаче.

Рассмотрим два примера.

Пример 1. Если вернуться к приведенной выше задаче о максимальной площади прямоугольника с заданным периметром $2p$, то на языке метода Лагранжа решение выглядело бы так. Поскольку $f(x, y) = xy$, а $F(x, y) = x + y - p$, то функция Лагранжа $L(x, y, \lambda) = xy - \lambda(x + y - p)$ и система (4) имеет вид

$$\frac{\partial L}{\partial x} = y - \lambda = 0,\ \frac{\partial L}{\partial y} = x - \lambda = 0,\ \frac{\partial L}{\partial \lambda} = -(x + y - p) = 0$$

Отсюда $x = y = \lambda$ и $2x = 2y = p$. Т.е. подозрение падает только на квадрат. По смыслу задачи максимум в ней должен существовать, значит он-то и найден (минимум площади равен нулю, если $x = 0$ или $y = 0$).

Пример 2. Найдем экстремумы симметрической квадратичной формы

$$f(x) = \sum_{i,j=1}^{n} a_{ij} x^i x^j \qquad (a_{ij} = a_{ji}) \tag{6}$$

на сфере

$$F(x) = \sum_{i=1}^{n} (x^i)^2 - 1 = 0 \tag{7}$$

Запишем функцию Лагранжа

$$L(x, \lambda) = \sum_{i,j=1}^{n} a_{ij} x^i x^j - \lambda \left(\sum_{i=1}^{n} (x^i)^2 - 1 \right)$$

данной задачи и с учетом того, что $a_{ij} = a_{ji}$, необходимые условия

$$\begin{cases} \frac{\partial L}{\partial x^i}(x, \lambda) = 2 \left(\sum_{j=1}^{n} a_{ij}(x^j) - \lambda x^i \right) = 0 \ (i = 1, \ldots, n) \\ \frac{\partial L}{\partial \lambda}(x, \lambda) = (\sum_{i=1}^{n} (x^i)^2 - 1) \end{cases} \tag{8}$$

экстремума.

Домножая первое уравнение на x^i и суммируя затем все первые соотношения, с учетом второго уравнения получим, что в точке экстремума должно быть выполнено равенство

$$\sum_{i,j=1}^{n} a_{ij} x^i x^j - \lambda = 0 \tag{9}$$

Систему (8) без последнего соотношения можно переписать в виде

$$\sum_{i,j=1}^{n} a_{ij} x^i x^j = \lambda x^i \qquad (i = 1, \ldots, n) \tag{10}$$

откуда видно, что λ — собственное значение линейного преобразования A, задаваемого матрицей (a_{ij}), а $x = (x^1, \ldots, x^n)$ — собственный вектор этого преобразования, отвечающий этому собственному значению.

Поскольку непрерывная на сфере (7) функция (6) обязана принимать в некоторой точке максимальное значение, система (8), а значит и система (10), должна иметь решение. Таким образом, мы попутно установили, что любая вещественная симметрическая матрица (a_{ij} имеет по крайней мере одно собственное значение. Это хорошо известный из линейной алгебры факт, связанный с построением базиса из собственных векторов симметрического преобразования.

Чтобы указать геометрический смысл собственного значения λ, заметим, что если $\lambda > 0$, то, переходя к координатам $t^i = x^i/\sqrt{\lambda}$, вместо (9) получим

$$\sum_{i,j=1}^{n} a_{ij} t^i t^j = 1 \tag{11}$$

а вместо (7)

$$\sum_{i=1}^{n} (t^i)^2 = \frac{1}{\lambda} \tag{12}$$

Но сумма в левой части (12) есть квадрат расстояния точки квадрики (11) до начала координат. Таким образом, если, например, соотношение (11) задает эллипсоид, то величина λ^{-1}, обратная к собственному значению λ, является квадратом величины одной из его полуосей.

Последнее наблюдение, кстати, показывает, что соотношения (8), необходимые для условного экстремума, еще не являются достаточными: ведь, например, в $\mathbb{R}^3$ эллипсоид кроме наибольшей и наименьшей полуосей может иметь и промежуточную по величине полуось, в любой окрестности конца которой есть как точки более близкие к началу координат, так и более далекие от него (проверьте).

Если бы мы искали экстремумы функции (12) на таком эллипсоиде (11), то среди подозрительных оказались бы и точки, связанные с промежуточными полуосями, которые, будучи критическими, не являются точками экстремума.

3. Геометрический смысл необходимого признака условного экстремума.

Пусть x_0 — точка экстремума (локального минимума или локального максимума) функции $f(x)$, ограниченной на гладкую поверхность S. Пусть $x = x(t)$ путь (закон движения некоторой точки) на поверхности S, такой, что $x_0 = x(0)$. Тогда функция $f(x(t))$ имеет экстремум при $t = 0$ и значит

$$\left.\frac{df(x(t))}{dt}\right|_{t=0} = f'(x_0)\dot{x}(0) = 0. \tag{13}$$

Вектор скорости $v = \dot{x}(0)$ лежит в плоскости $T_{x_0}S$, касательной к S в точке x_0 и, как мы знаем, вся плоскость (касательное пространство) $T_{x_0}S$ состоит из таких векторов.

Если считать, что $f'(x_0) \neq 0$, т.е. что x_0 не является критической и тем самым подозрительной точкой уже самой функции f (а не $f|_S$), то соотношение

$$f'(x_0)(x - x_0) = 0 \tag{14}$$

можно считать уравнением плоскости $T_{x_0}N$, касательной к поверхности $N = \{x | f(x) = f(x_0)\}$ уровня функции f в точке x_0.

Сопоставляя (13) и (14), заключаем, что если x_0 — точка экстремума функции $f|_S$ (т.е. точка условного экстремума функции $f(x)$ при $x \in S$), то в этой точке необходимо выполняется следующее включение:

$$T_{x_0}S \subset T_{x_0}N \tag{15}$$

Пример 3. Найдем экстремумы функции $f(x, y) = xy$ на окружности $x^2 + y^2 = 1$.

Если нарисовать (см. рис. 1) линии уровня функции f, то в соответствии со сказанным, экстремумы $f|_S$ могут быть только среди отмеченных четырех точек взаимного касания окружности и линий уровня функции f. Остается сравнить значения функции в этих точках и убедиться, что в точках $\pm(1, 1)$ $f|_S$ имеет максимальное значение 1, а в двух других — минимальное -1.

Рис.1.

Необходимый признак (15) условного экстремума проиллюстрирован на рисунках 2(a, b, c).

Первый из них показывает, что если поверхность S пересекает поверхность N уровня функции f, трасверсально (не касательно), то (при условии что $\operatorname{grad} f(x_0) \neq 0$) она пересечет и соседние уровни как больших, так и меньших значений функции f. Значит в этом случае x_0 заведомо не является точкой экстремума для $f|_S$.

Рисунок 2b показывает, как S, касаясь N, может локально оказаться по одну сторону N и функция $f|_S$ в этом случае будет иметь локальный экстремум.

Рис.2.

Рисунок 2с объясняет, почему полученное необходимое условие не является достаточным для наличия экстремума. Он соответствует отысканию экстремума функции $f(x, y) = y$ на кривой S, заданной уравнением $F(x, y) = x^3 - y = 0$, при этом и $\operatorname{grad} f(0, 0) \neq 0$, и $\operatorname{grad} F(0, 0) \neq 0$. Кстати, функция $y = x^3$ служит стандартным примером того, что необходимое условие внутреннего (свободного) экстремума не является достаточным. По существу здесь обыгран тот же пример.

Рис.3.

Наконец заметим, что все эти картинки изображали регулярную (локально распрямляемую) картину поверхностей (линий) уровня рассматриваемых функций, которая, как мы знаем, действительно имеет место в окрестности любой не критической (регулярной) точки, т.е. точки, где производная имеет максимально возможный ранг. Если же x_0 — критическая точка f или F, то формулы (14), (16), определявшие в регулярном случае касательные пространства, перестают действовать, а взаимное расположение S и N может быть совсем иным. Рисунок 3, например, отвечает случаю $f(x, y) = y$, $F(x, y) = x^2 - y^3$. Никакого касания S и N в начале координат здесь нет, а экстремум (строгий минимум $f|_S$), очевидно, имеется.

Замечание. В соотношении (15) включение, конечно, может быть и строгим: ведь $T_{x_0}N$ — всегда гиперплоскость, а $T_{x_0}S$ имеет размерность S, так, что если S, например, кривая, то $\dim T_{x_0}S = 1$.

Если поверхность S задана в виде системы связей (2), т.е. в виде векторного равенства $F(x) = 0$, то T_{x_0} определяется, подобно (14), уравнением (системой)

$$F'(x_0)(x - x_0) = 0 \tag{16}$$

В координатной записи, соответствующей формулам (1) и (2), уравнения (14) и (16) имеют вид

$$\frac{\partial f}{\partial x^1}(x_0)(x^1 - x_0^1) + \cdots + \frac{\partial f}{\partial x^m}(x_0)(x^m - x_0^m) = 0 \tag{14'}$$

и

$$\begin{cases} \frac{\partial F_1}{\partial x^1}(x_0)(x^1 - x_0^1) + \cdots + \frac{\partial F_1}{\partial x^m}(x_0)(x^m - x_0^m) = 0, \\ \ldots\ldots \qquad \ldots\ldots \\ \frac{\partial F_k}{\partial x^1}(x_0)(x^1 - x_0^1) + \cdots + \frac{\partial F_k}{\partial x^m}(x_0)(x^m - x_0^m) = 0. \end{cases} \tag{16'}$$

Соотношение (15) означает, что каждое решение системы (16′) является решением уравнения (14′), т.е. равенство (14′) является следствием системы (16′), или иначе, вектор его коэффициентов является линейной комбинацией векторов-строк матрицы $F'(x_0)$. Т.е.

$$\operatorname{grad} f(x_0) = \lambda_1 \operatorname{grad} F_1(x_0) + \cdots + \lambda_k \operatorname{grad} F_k(x_0) \tag{15'}$$

Так, что соотношение (15) в рассматриваемом случае равносильно соотношению (15′) или, записанному в более компактной форме, равенству

$$f'(x_0) = \lambda F'(x_0) \tag{15''}$$

Мы снова, но уже из других соображений, получили необходимый признак условного экстремума. Чтобы помнить, что еще и $F(x_0) = 0$ полезно, как и в (5′), использовать функцию Лагранжа

$$L(x, \lambda) = f(x) - \lambda F(x) \tag{17}$$

4. Физический (механический) смысл множителей Лагранжа.

Поскольку мы искали точку экстремума x_0, то мы пока не очень обращали внимание на вспомогательные параметры $\lambda = (\lambda_1, \ldots, \lambda_k)$, участвующие в функции Лагранжа (5′), (17) и в необходимом признаке (15′), (15″) условного экстремума. Им, однако, можно дать следующую физическую интерпретацию.

Пусть материальная частица, имеющая возможность перемещаться, не покидая идеально гладкой поверхности S, находится в потенциальном поле сил $-\operatorname{grad} U$ с потенциалом U. Ищем точки устойчивого равновесия частицы на этой поверхности, т.е. точки локального минимума функции $U|_S$.

Предположим поверхность S задана в форме одного уравнения $F(x^1, \dots, x^m) = 0$, причем $F'(x) = \operatorname{grad} F \neq 0$ на S.

Если частица находится в равновесии, то сумма действующих на нее сил равна нулю. У нас этих сил всего две: действие поля $-\operatorname{grad} U(x)$ (например, гравитационного поля в виде силы тяжести) и реакция связи, которая ввиду предположенной идеальной гладкости S, направлена только по нормали к поверхности. Значит в точке равновесия должно быть выполнено равенство

$$\operatorname{grad} U(x_0) = \lambda \operatorname{grad} F(x_0) \tag{18}$$

причем, если $|\operatorname{grad} F(x_0)| = 1$, (чего всегда можно добиться, поделив F на $|\operatorname{grad} F(x_0)|$), то величина λ как раз дает величину силы реакции связи. Мы в третий раз независимо пришли к необходимому признаку условного экстремума, с той лишь разницей, что теперь вместо f фигурирует функция U.

Если движение происходит не по поверхности, а например, по кривой (бусинка на проволочке), то вместо одного равенства $F(x^1, \dots, x^m) = 0$ появится система (2) равенств и по тем же соображениям о нулевой равнодействующей в положении равновесия, получим вместо (18) аналогичное (15$'$) равенство

$$\operatorname{grad} U(x_0) = \lambda_1 \operatorname{grad} F_1(x_0) + \dots + \lambda_k \operatorname{grad} F_k(x_0)$$

в котором $\lambda_1, \dots, \lambda_k$ укажут величину реакций соответствующих связей $F_1, \dots, F_k$, при условии, что $|\operatorname{grad} F_i(x_0)| = 1$ на поверхностях S_i $(i = 1, \dots, k)$, заданных уравнениями $F_i(x) = 0$ $(i = 1, \dots, k)$.

5. Алгебраический аспект и обобщающая формулировка метода Лагранжа.

Рассуждения, которые быстро привели нас к геометрическому признаку (15) условного экстремума в равной мере относятся к гладкой функции $f : X \to \mathbb{R}$, определенной, например, в любом векторном нормированном пространстве X (или его части). В равенстве $F(x) = 0$, выделяющем интересующую нас часть S области определения f, можно считать F гладким отображением $F : X \to Y$ пространства X (или подобласти X) в пространство Y (так что в равенстве $F(x) = 0$ ноль — это ноль в Y).

Если уравнения (14) и (16) задают соответствующие касательные пространства $T_{x_0}N$ и $T_{x_0}S$, то, рассматривая их как ядра линейных отображений $f'(x_0) : X \to \mathbb{R}$ и $F'(x_0) : X \to Y$, включение (15) можно переписать в виде

$$\ker F'(x_0) \subset \ker f'(x_0) \tag{15'''}$$

Но справедлива следующая простая алгебраическая лемма.

Лемма. *Если линейные отображения $A : X \to Z$ и $B : \to Y$ таковы, что $\ker B \subset \ker A$, то существует линейное отображение $\lambda : Y \to Z$ такое, что $A = \lambda B$ (см. рис. 4).*

$$\begin{tikzcd} X \arrow[rr, "B"] \arrow[dr, "A"'] & & Y \arrow[dl, "\lambda"] \\ & Z & \end{tikzcd}$$

Рис.4.

Доказательство. Если $y = Bx$, то положим $\lambda y = Ax$ (ведь мы желаем равенства $A = \lambda B$). Это определение корректно, т.к. если $y = Bx_1 = Bx_2$, то $B(x_1 - x_2) = 0$, а тогда по условию $A(x_1 - x_2) = 0$, т.е. $Ax_1 = Ax_2$. Итак, отображение $\lambda : Y' \to$ определено на $B(X) = Y' \subset Y$. Разложив Y в прямую сумму $Y' \oplus Y''$ подпространств, для любого вектора y, имея однозначное разложение $y = y' + y''$, положим $\lambda(y) = \lambda(y')$. Мы построили отображение $\lambda : Y \to Z$ и, очевидно, $A = \lambda B$.

Возвращаясь к соотношению $(15''')$, можно теперь сказать, что существует линейное отображение $\lambda : Y \to \mathbb{R}$ такое, что выполнено равенство $(15'')$.

Таким образом, функция Лагранжа (17), метод Лагранжа и необходимое условие экстремума (15), $(15'')$, $(15''')$ остаются в силе и в общей ситуации. В том случае, когда $Y = \mathbb{R}^k$, линейная функция $\lambda : \mathbb{R}^k \to \mathbb{R}$ задается набором числовых коэффициентов $(\lambda_1, \ldots, \lambda_k)$ и равенство $(15'')$ конкретизируется в равенство $(15')$.

Замечание. В проведенных выше исследованиях необходимого признака условного экстремума, в записи (14), (16) касательных пространств, мы считали $f'(x_0)$ и $F'(x_0)$ отображениями максимально возможного ранга. В первом случае это означало, что $f'(x_0) \neq 0$. Если $f'(x_0) = 0$, то точка x_0 подозрительна не только в смысле условного, но и в смысле свободного экстремума функции f и поэтому не содержит элемент новизны теории, хотя при вычислениях должна быть учтена и функция Лагранжа (17) помимо нашей воли об этом, очевидно, позаботилась. Мы требовали также максимальности ранга $F'(x_0)$, подразумевая с одной стороны, что лишние (зависимые) условия всегда можно отбросить, а с другой, что мы исследуем типичную неособую точку или, как говорят, точку общего положения. Однако экстремум может оказаться в некоторой критической точке (где rang $F'(x_0)$ не максимально возможный). Это изображено на рисунке 3. Конечно, такие точки можно рассмотреть особо. Удобнее, однако, включить их в общее рассмотрение, если функцию Лагранжа писать в виде

$$L(x, \lambda) = \lambda_0 f(x) + \lambda F(x) \tag{17'}$$

Если $\lambda_0 \neq 0$, то эта функция приводит к тем же условиям (уравнениям) для точек экстремума, что и функция (17).

Но теперь в рассмотрение включен и случай вырождения $F'(x_0)$, которому будет отвечать $\lambda_0 = 0$.

Задача. Проведите исследование соответствующего рисунку 3 примера, используя функцию Лагранжа (17′).

5. Достаточный признак условного экстремума.

Дадим в заключение один достаточный признак условного экстремума.

Утверждение. *Пусть ищется экстремум функции $f \in C^{(2)}(D, \mathbb{R})$ на поверхности $S \subset D$, заданной уравнением $F(x) = 0$ (система функций класса $C^{(2)}(D, \mathbb{R})$).*

Пусть $L(x, \lambda) = f(x) - \lambda F(x)$ — функция Лагранжа задачи и пусть точка $x_0 \in S$ и параметр λ_0 выбраны в соответствии с необходимым условием экстремума

$$L'(x_0, \lambda_0) = 0$$

В окрестности x_0 рассмотрим функцию $L_0(x) = f(x) - \lambda_0 F(x)$ переменной x.

Первый дифференциал dL_0 этой функции равен нулю при $x = x_0$.

Если при этом ее второй дифференциал — квадратичная форма d^2L в точке $x = x_0$ положительно или отрицательно определена на векторах пространства $T_{x_0}S$, то функция $f|_S$ имеет в x_0 строгий локальный минимум или строгий локальный максимум соответственно.

Если же эта форма может на T_{x_0} принимать значения разных знаков, то x_0 не является точкой экстремума функции $f|_S$.

Доказательство. Заметим прежде всего, что $L(x, \lambda) = f(x)$ при $x \in S$, поэтому $f(x) - f(x_0) = L_0(x) - L_0(x_0)$ при $x \in S$. Учитывая это, изучим поведение последней разности в окрестности точки x_0 на поверхности S. По формуле Тейлора

$$L_0(x) - L_0(x_0) = \frac{1}{2}d^2L_0(x_0)(x - x_0)^2 + o(|x - x_0|^2) \tag{19}$$

Если x' — проекция точки $x \in S$ на пространство $T_{x_0}S$ (направление проектирования соответствующей размерности при этом не существенно, но фиксировано), то по определению касательного пространства при $x \to x_0$ имеем $x - x_0 = (x' - x_)) + o(|x' - x_0|)$.

Значит в правой части равенства (19) вместо x можно поставить x'. Тогда вектор $x' - x_0$ окажется в пространстве $T_{x_0}S$, и мы приходим к стандартному анализу положительности или отрицательности правой части формулы (19) уже не на S, а на векторном пространстве $T_{x_0}S$. Но именно такой анализ мы уже проделали, доказывая достаточные условия наличия или отсутствия обычного внутреннего экстремума функции.

Контрольный вопрос:

1. Верно ли, что если в точке $x_0 \in S$ выполнен необходимый признак условного экстремума функции f на S, то $df(x_0)v = 0$ для любого вектора $v \in T_{x_0}S$? (Конечно!)
2. Можно ли тогда в равенстве (19) вместо L_0 всюду поставить саму функцию f? (Почему?)

Во многих аспектах поучительны следующие простые в вычислительном и геометрическом плане примеры, на которых можно увидеть и прощупать как необходимые, так и достаточные условия экстремума, в том числе роль параметров и неформальную роль самой функции Лагранжа.

Задача.

1. Найдите экстремумы функции $f(x, y) = x^2+y^2$ на эллипсе $F(x, y) = x^2/a^2 + y^2/b^2 - 1 = 0 \quad (0 < a < b)$.
2. Обратите внимание, здесь есть и минимумы и максимумы, хотя форма d^2f во всех точках положительно определена.
3. Выполните предыдущие задания применительно к функции $f(x, y, z) = x^2+y^2+z^2$ на эллипсоиде $F(x, y, z) = x^2/a^2+y^2/b^2+z^2/c^2-1 = 0 \quad (0 < a < b < c)$ и убедитесь еще раз, что приведенные необходимые условия экстремума, вообще говоря, не являются достаточными.