О р д е н а   Л е н и н а
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В. Келдыша
А к а д е м и и   н а у к   С С С Р

Е.А. Морозова, Н.Н. Ченцов

ИНВАРИАНТНЫЕ ЛИНЕЙНЫЕ СВЯЗНОСТИ НА СОВОКУПНОСТЯХ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ

Препринт № 167 за 1988г.

Москва

АКАДЕМИЯ НАУК СССР

ОРДЕНА ЛЕНИНА ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ

ИМЕНИ М.В. КЕЛДЫША

Е.А.Морозова, Н.Н.Ченцов

ИНВАРИАНТНЫЕ ЛИНЕЙНЫЕ СВЯЗНОСТИ НА СОВОКУПНОСТЯХ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ

Москва, 1988 г.

Морозова Е.А., Ченцов Н.Н. Инвариантные линейные связности на совокупностях распределений вероятностей. Препринт // М.: Ин-т прикл. матем. им. М.В.Келдыша АН СССР. 1988. № 167.

Авторами показано, см. РЖ Мат, 1973, 4В250К, что все линейные связности на совокупностях распределений вероятностей $Cap(\Omega, \mathcal{A})$, инвариантные относительно категории марковских отображений, составляют однопараметрическое семейство ${}^{\gamma}\nabla$, $\gamma \in \mathbb{R}$. Связность с $\gamma = \frac{1}{2}$ отвечает единственной (с точностью до постоянного множителя) инвариантной в категории римановой метрике. В настоящей работе для всех $\gamma \in \mathbb{R}$ выписываются явные выражения геодезических, проходящих через две "точки" или через "точку" по данному направлению; показано, что у всех инвариантных связностей тензор кручения тождественно равен нулю, а тензоры кривизны Римана-Кристоффеля пропорциональны,

$${}^{\gamma}R^{k}_{\ell ij} = 4\gamma(\gamma-1)\cdot {}^{\frac{1}{2}}R^{k}_{\ell ij}$$

. Эти результаты обобщают данные Амари.

В категории САР существует дополнительная операция тензорного умножения объектов. Мы доказываем, что единственно инвариантной связностью, наследственной относительно возведения в степень, является только логарифмическая плоская связность, отвечающая $\gamma = 0$, геодезическими линиями и вполне геодезическими поверхностями в которой являются экспонентные семейства распределений в канонической параметризации.

Morozova E.A., Čencov N.N. Invariant linear connections on probability distribution collections. Preprint // M.: Keldysh Inst. Appl. Mathem., the USSR Ac. Sci. 1988. N 167.

We had proved that the invariant (under Markov map category) linear connections on probability distribution collections $Cap(\Omega, \mathcal{A})$ form one parameter family ${}^{\gamma}\nabla$, $\gamma \in \mathbb{R}$, see Math. Rev. 49 8140; 83 g : 62004. The connection ${}^{1/2}\nabla$ turned out to be generated by the unique (up to a constant factor) invariant Riemann metric. In the paper the simple expressions of geodesics passing through two "points" or going from a "point" in a given "direction" are written out; for all γ the torsions are shown to be zero identically, and the curvature tensors of Riemann-Christoffel are proved to be proportional, i.e. ${}^{\gamma}R^{k}_{lij} = 4\gamma(\gamma-1) \cdot {}^{1/2}R^{k}_{lij}$. These results extend those given by Amari, see Math. Rev. 86 m : 62053.

In the category CAP the supplementary operation of (tensor) multiplication of objects $Cap(\Omega, \mathcal{A})$ is of a great moment, comp. Math. Rev. 83 b : 62001. Now we prove that the logarithmic plane connection ${}^{0}\nabla$ is the only invariant one to be inheritable under raising to the power; in ${}^{0}\nabla$ the geodesic lines and completely geodesic surfaces are well known to be exponent distribution families in canonical parametrization.

Введение

В монографии [1], § 12, см. также [2], было показано, что все линейные связности на совокупностях распределений вероятностей $\mathrm{Cap}(\Omega, \mathcal{A})$, инвариантные относительно категории марковских отображений (статистических решающих правил), составляют однопараметрическое семейство ${}^{\gamma}\nabla$, $\gamma \in \mathbb{R}$. В последующей лекции по основам статистической теории, см. [3], было отмечено, что в категории CAP следует учитывать дополнительную операцию тензорного умножения объектов. Поэтому, основной интерес должны представлять категорные эквиварианты, наследственные относительно возведения в степень. В этой работе мы показываем, что среди всех связностей ${}^{\gamma}\nabla$ последнему условию удовлетворяет только подробно изученная нами в [1] логарифмическая плоская связность ${}^{0}\nabla$, геодезическими линиями и вполне геодезическими поверхностями в которой являются экспонентные семейства распределений в канонической параметризации. Нами также даются явные выражения для геодезических в связностях ${}^{\gamma}\nabla$, ср. [5], показывается, что их кручение равно нулю, а тензоры кривизны Римана-Кристоффеля - пропорциональны.

§ 1. Способы задания векторных полей на объектах

Каждое распределение вероятностей P на алгебре $\mathcal{A}_n$ всех подмножеств множества $\Omega_n = \{\omega_1, \ldots, \omega_n\}$ всех исходов ω задается вектором-строкой $\vec{p} = \langle\!\langle p_1, \ldots, p_n \rangle\!\rangle$ вероятностей атомов, $\forall j$, $p_j = \mathrm{P}\{\omega_j\}$. Для распределения вероятностей:

$$\forall j \;, \quad p_j \geqslant 0 \;; \qquad\qquad \sum_{j=1}^{n} p_j = 1 \;. \qquad (1.1)$$

Обозначим через $\vec{e}(k)$, $k = 1, \ldots, n$, вектор распределения, сосредоточенного на ω_k :

$$e_j(k) = \delta_j^k \quad . \tag{I.2}$$

Тогда в каждой внутренней точке $\vec{p}$ симплекса (I.I) возникнет переполненный (на один вектор) базис

$$\vec{e}(k) - \vec{p} \;, \qquad k = 1,\ldots,n \;; \tag{I.3}$$

и связанные с ним реперы

$$p_k \vec{e}(k) - p_k \vec{p} \quad , \qquad k = 1,\ldots,n \;; \tag{I.4}$$

$$(1-p_k)^{-1}\cdot\left[\vec{e}(k) - \vec{p}\,\right] = \vec{e}(k) - \vec{q}(k) \;, \tag{I.5}$$

где вектор $\vec{q}(k)$ – вектор вероятностей атомов для условного распределения $P\{\cdot \mid \Omega - \omega_k\}$, лежащий в k-ой грани симплекса. Для каждого фиксированного вектора $\vec{q}(k)$ чевиана симплекса

$$\vec{p}(t) = t\,\vec{e}(k) + (1-t)\vec{q}(k) = \vec{q}(k) + t\left[\vec{e}(k) - \vec{q}(k)\right] \tag{I.6}$$

при $0 \leqslant t \leqslant 1$ задает семейство распределений с одним и тем же вектором условных вероятностей $\vec{q}(k)$. Расслоение симплекса на чевианы, исходящие из $\vec{e}(k)$, задает внутри симплекса поле касательных векторов – дифференцирование по натуральному параметру t:

$$(Z_k f)_{\vec{p}} = \frac{d}{dt} f(\vec{p}(t))\Big|_{\vec{p}(t)=\vec{p}} \quad . \tag{I.7}$$

Будем считать, что гладкая функция $f(\cdot)$ задана на всем положительном ортанте

$$\vec{z} \;:\; \forall j \;,\; z_j \geqslant 0 \;, \tag{I.8}$$

а не только на симплексе (I.I); например, продолжена с (I.I) на (I.8) по однородности:

$$f(h p_1,\ldots, h p_n) = h\cdot f(p_1,\ldots,p_n)\,, \quad \forall h>0. \tag{I.9}$$

Тогда для дифференцирования вдоль любой гладкой кривой $\vec{p}(\theta)$ справедливо

$$\frac{d}{d\theta} f(\vec{p}(\theta)) = \frac{\partial f}{\partial p_1} \frac{dp_1}{d\theta} + \ldots + \frac{\partial f}{\partial p_n} \frac{dp_n}{d\theta} \quad . \qquad (\text{I}.10)$$

Эту связь мы будем символически записывать

$$\frac{d}{d\theta} \leftrightarrow \ll p_1'(\theta), \ldots, p_n'(\theta) \gg = \vec{p}\,'(\theta) . \qquad (\text{I}.11)$$

Тогда для дифференцирования Z_k по натуральному параметру t вдоль чевиан k-ого семейства ;

$$Z_k \leftrightarrow \vec{e}(k) - \vec{q}(k) = (1 - p_k)^{-1} \left[\vec{e}(k) - \vec{p} \right] . \qquad (\text{I}.12)$$

Перепараметризуем чевиану так, чтобы

$$\frac{d}{ds} = t(1-t)\frac{d}{dt} \quad , \quad ds = \frac{dt}{t} + \frac{dt}{1-t} \quad . \qquad (\text{I}.13)$$

Векторные поля X_k дифференцирований по параметру s ,

$$X_k = \frac{d}{ds_k} \leftrightarrow p_k \vec{e}_k - p_k \vec{p} \ , \qquad (\text{I}.14)$$

оказываются более удобными для описания линейных связностей. В силу связи (I.14)

$$0 = \vec{p} - \vec{p} = \sum_{k=1}^{n} \left[p_k \vec{e}(k) - p_k \vec{p} \right] \leftrightarrow \sum_{k=1}^{n} X_k ,$$

откуда

$$\sum_{k=1}^{n} X_k = 0 . \qquad (\text{I}.15)$$

Лемма I.I.

Любой касательный вектор $V = \frac{d}{d\theta}$

$$(Vf)_{\vec{p}} = \sum_{k=1}^{n} v_k(\vec{p}) \frac{\partial f}{\partial p_k}\Big|_{\vec{p}} \quad , \quad \sum_{k=1}^{n} v_k = 0 , \qquad (\text{I}.16)$$

удобно раскладывать по простейшему переполненному базису $(Y_j)_{\vec{p}}$ $\leftrightarrow \vec{e}(j) - \vec{p}$;

$$V = \sum_{j=1}^{n} v_j Y_j \quad . \qquad (\text{I}.17)$$

Доказательство. В самом деле

$$\sum_{j=1}^{n} v_j Y_j \leftrightarrow \sum_{j=1}^{n} v_j \left[\vec{e}(j) - \vec{p}\,\right] = \sum_{j=1}^{n} v_j \vec{e}(j) - \vec{p}\sum_{j=1}^{n} v_j =$$

$$= \sum_{j=1}^{n} v_j \vec{e}(j) = \langle\!\langle v_1, \ldots, v_n \rangle\!\rangle \leftrightarrow V .$$

В остальных реперах одно из представлений можно получить подстановкой. Например,

$$V = \sum_{j=1}^{n} \frac{v_j}{p_j} X_j = \sum_{j=1}^{n} \xi_j^V(\vec{p}) X_j \; ; \; \sum_{j=1}^{n} \xi_j^V(\vec{p}) p_j = 0. \quad (\text{I}.18)$$

Ввиду (I.15) общий вид разложения

$$V = \sum_{j=1}^{n} \left(\frac{v_j}{p_j} - c \right) X_j \; , \quad \forall c \in \mathbb{R} . \quad (\text{I}.19)$$

Построим теперь двумерную поверхность, расслаивающуюся одновременно на чевианы j -го и k -го семейств. Для упрощения обозначений возьмем $j = 1$, $k = 2$.

<u>Лемма I.2</u>.

Семейство $P_{\vec{s}}(\cdot)$, $\vec{s} = \langle\!\langle s_1 , s_2 \rangle\!\rangle$, описываемое семейством векторов $\vec{p}(s_1, s_2) \in \operatorname{Int} \operatorname{Cap}(\Omega_n, \mathcal{A}_n)$,

$$\begin{aligned} p_1(s_1, s_2) &= p_1(0,0) \exp[s_1 - \Psi(s_1, s_2)] , \\ p_2(s_1, s_2) &= p_2(0,0) \exp[s_2 - \Psi(s_1, s_2)] , \\ p_i(s_1, s_2) &= p_i(0,0) \exp[-\Psi(s_1, s_2)], \quad i \geqslant 3 ; \end{aligned} \quad (\text{I}.20)$$

$$\Psi(s_1, s_2) = \ln\left[p_1(0,0)\exp s_1 + p_2(0,0)\exp s_2 + \sum_{i \geqslant 3} p_i(0,0)\right],$$

при фиксированном значении $s_2 = \sigma$ задает чевиану первого семейства в канонической параметризации, связанной с натуральной координатой t_1 соотношением (I.13). Аналогичная конструкция задает чевианы второго семейства.

<u>Следствие</u>. Коммутатор базисных полей

$$[X_j , X_k] = X_j X_k - X_k X_j = 0 , \quad \forall j, k . \quad (\text{I}.21)$$

Доказательство. Обозначим $t = t(s_1, \sigma) = p_1(s_1, \sigma)$, и введем вектор $q(1) = [1 - t(0,\sigma)]^{-1}\{\vec{p}(0,\sigma) - t(0,\sigma)\vec{e}(1)\}$, $q_1(1)=0$. Легко проверить по-компонентно, что

$$\forall s_1\ ,\quad [1 - t(s_1,\sigma)]^{-1}\{\vec{p}(s_1,\sigma) - t(s_1,\sigma)\vec{e}(1)\} = \vec{q}(1). \tag{1.22}$$

Таким образом, обратно $\vec{p}(s_1,\sigma)$ образуют чевиану

$$p(s_1,\sigma) = t(s_1,\sigma)\vec{e}(1) + [1 - t(s_1,\sigma)]\,\vec{q}(1)\ . \tag{1.23}$$

Найдем связь между дифференцированием по s_1 и по t . Так как

$$\frac{d\Psi(s_1,\sigma)}{d s_1} = p_1\ , \tag{1.24}$$

то, дифференцируя по-компонентно, находим $\vec{p}'_{s_1} = p_1\vec{e}_1 - p_1\vec{p}$, откуда $\frac{\partial}{\partial s_1} = X_1$ в силу (1.13) и (1.14). Аналогично, $\frac{\partial}{\partial s_2} = X_2$. На поверхности (1.20) $\frac{\partial^2\vec{p}}{\partial s_1\partial s_2} = \frac{\partial^2\vec{p}}{\partial s_2\partial s_1}$, т.е. $X_1X_2 = X_2X_1$, ч.т.д.

§ 2. Геодезическая параметризация чевиан

Любая пара взаимно дизъюнктных распределений вероятностей конгруентна любой другой такой паре, см. следствие из теоремы 6.2 в [1] . Поэтому, любая чевиана (1.6) конгруентна простейшей совокупности Cap_2:

$$\{\langle\langle p_1, p_2\rangle\rangle : p_1 = t\ ,\ p_2 = 1-t\ ,\ 0\leqslant t\leqslant 1\} = \mathrm{Cap}_2\ , \tag{2.1}$$

- отрезку с концами $\vec{e}(1)$ и $\vec{e}(2)$. Этот отрезок обязан быть геодезической в любой инвариантной линейной связности. Имеем по (1.13) и (1.14)

$$n=2\ ,\quad X_1 = -X_2 = X = t(1-t)\frac{d}{dt}\ . \tag{2.2}$$

Согласно теореме 12.2 из [1] для связности ${}^{\gamma}\nabla$ выполнены соотношения

$${}^{\gamma}\nabla_{X_k}X_j = -\gamma\,[p_j X_k + p_k X_j]\ ,\quad j\neq k\ , \tag{2.3}$$

$$^{\gamma}\nabla_{X_j} X_j = \gamma\,(1-2p_j)\,X_j \quad . \qquad (2.4)$$

Для геодезической (2.I) оба эти соотношения ввиду (2.2) эквивалентны правилу переноса

$$^{\gamma}\nabla_X X = \gamma\,(1-2t)\,X \; . \qquad (2.5)$$

Теорема 2.I. Поле параллельных вдоль геодезической (2.I) касательных векторов $^{\gamma}V$ в инвариантной линейной связности $^{\gamma}\nabla$ задает поле дифференцирований

$$\frac{d}{d\theta} = V = \left[t(1-t)\right]^{-\gamma} X = \left[t(1-t)\right]^{-\gamma+1} \frac{d}{dt} \, . \qquad (2.6)$$

Доказательство. Вычислим $^{\gamma}\nabla_V V$, используя (2.5), (2.2) и основные свойства символа ∇ , см. [4] гл. I § 4 или [I] § 3 ,

$$\nabla_V V = \nabla_{[\;\;]^{-\gamma}X} [\;\;]^{-\gamma} X = [\;\;]^{-\gamma} \nabla_X [\;\;]^{-\gamma} X =$$

$$= [\;\;]^{-\gamma} \left(X [\;\;]^{-\gamma} \right) X + [\;\;]^{-2\gamma} \nabla_X X =$$

$$= -[t(1-t)]^{-\gamma}\, t(1-t)\, \gamma\, [t(1-t)]^{-\gamma-1} (1-2t) X +$$

$$+ [t(1-t)]^{-2\gamma}\, \gamma\, (1-2t)\, X = 0 \, .$$

Из доказанной теоремы нетрудно получить геодезические параметризации при разных γ .

При $\gamma = 1$: $^{1}V = \dfrac{d}{dt}$, $d\overset{1}{\theta} = dt$, $\overset{1}{\theta} = t - t_0$.

При $\gamma = 0$: $^{0}V = X$, $\overset{0}{\theta} = s$,

$$ds = \frac{dt}{t} + \frac{dt}{1-t} \; ;$$

$$t = \frac{\exp(s/2)}{\exp(s/2) + \exp(-s/2)} \; , \quad s = \ln t - \ln(1-t) . \qquad (2.7)$$

Следствие I. При $\gamma = 0$ геодезический параметр s является каноническим параметром экспонентного семейства $\{\langle\!\langle p_1(s), p_2(s) \rangle\!\rangle\}$

$$p_k(s) = \frac{\exp[(-1)^{k-1} s/2]}{\exp(s/2) + \exp(-s/2)}, \quad -\infty \leq s \leq \infty . \qquad (2.8)$$

А его натуральным экспонентным параметром является t , т.е. геодезический параметр, отвечающий $\gamma = 1$.

При $\gamma = 1/2$ линейная связность отвечает римановой метрике, см. [1] , § 12.

Следствие 2. При $\gamma = 1/2$ геодезическим параметром является риманова длина дуги σ , отвечающая Фишеровой квадратичной форме:

$$(d\sigma)^2 = \frac{(dt)^2}{t} + \frac{[d(1-t)]^2}{1-t} = 4(d\sqrt{t})^2 + 4(d\sqrt{1-t})^2 ; \qquad (2.9)$$

$$d\sigma = [t(1-t)]^{-1/2} dt . \qquad (2.10)$$

С точностью до выбора начала отсчета

$$\sigma - \sigma_0 = \arccos(1-2t) = 2 \arcsin\sqrt{t} = 2 \arccos\sqrt{1-t} . \qquad (2.11)$$

Следствие 3. Для γ общего вида геодезический параметр выражается через t и σ квадратурами

$$d\theta = [t(1-t)]^{\gamma-1} dt = 2^{-2(\gamma-1)} \left[\sin\frac{\sigma}{2}\right]^{\gamma} d\sigma =$$

$$= 2^{-2\gamma+1} [\sin\varphi]^{\gamma} d\varphi \qquad (2.12)$$

где $\varphi = \sigma/2$, $\sin\varphi = \sqrt{t}$, $\cos\varphi = \sqrt{1-t}$, $0 \leq \varphi \leq \frac{\pi}{2}$.

Как отмечалось в начале этого параграфа, каждая чевиана конгруентна совокупности Cap_2 . Следовательно, она является геодезической линией при любом γ , и все соотношения (2.7)-(2.12) между геодезическими параметрами сохраняют силу и для нее.

§ 3. Квадрат совокупности Cap_2

Если инвариантная линейная связность ${}^{\gamma}\nabla$ наследственна в тензорных степенях, то семейство тензорных квадратов распределений вероятностей совокупности Cap_2 должна быть геодезической линией (с тем же геодезическим параметром) в этой связности ${}^{\gamma}\nabla$, поскольку геодезической линией является сама Cap_2. Выясним сейчас, при каких γ будет так.

Пусть даны два экземпляра множества $\Omega_2 = \{\omega_1, \omega_2\}$: Ω_2' и Ω_2''.

$$\Omega_2' \times \Omega_2'' = \{\omega_1'\omega_1'', \omega_1'\omega_2'', \omega_2'\omega_1'', \omega_2'\omega_2''\}.$$

Лексикографически перенумеровав исходы единой нумерацией, получаем $\Omega_2' \times \Omega_2'' = \Omega_4 = \{\omega_1, \omega_2, \omega_3, \omega_4\}$.

Если $\vec{p} = \langle\!\langle p_1, p_2 \rangle\!\rangle$ — распределение на $(\Omega_2', \mathcal{A}_2')$, $\vec{q} = \langle\!\langle q_1, q_2 \rangle\!\rangle$ — на $(\Omega_2'', \mathcal{A}_2'')$, то их произведение на $(\Omega_4, \mathcal{A}_4)$ задается вектором

$$\vec{p} \otimes \vec{q} = \langle\!\langle p_1 q_1, p_1 q_2, p_2 q_1, p_2 q_2 \rangle\!\rangle. \tag{3.1}$$

Отсюда, полагая $p_i = q_i$, $i = 1, 2$, квадрат совокупности Cap_2, взятой в натуральной параметризации, будет описываться кривой

$$\vec{r}(t) = \langle\!\langle t^2, t(1-t), t(1-t), (1-t)^2 \rangle\!\rangle, \quad 0 \leqslant t \leqslant 1. \tag{3.2}$$

Поэтому, семейство касательных векторов относительно исходной натуральной $(\gamma = 1)$ параметризации

$$\dot{\vec{r}}(t) = \langle\!\langle 2t, 1-2t, 1-2t, -2+2t \rangle\!\rangle. \tag{3.3}$$

Относительно любой другой исходной γ- геодезической параметризации в силу связи (2.12)

$$\vec{r}_\theta{}'(t(\theta)) = \dot{\vec{r}}(t) \cdot t_\theta' = [t(1-t)]^{-\gamma+1} \dot{\vec{r}}(t). \tag{3.4}$$

На кривой $\vec{r}(t(\theta))$ дифференцирование по параметру $d/d\theta$ задает семейство касательных векторов $(W)_{\vec{r}} \leftrightarrow [t(1-t)]^{-\gamma+1} \dot{\vec{r}}(t)$, т.е.

$$\langle\!\langle w_1, w_2, w_3, w_4 \rangle\!\rangle = [t(1-t)]^{-\gamma+1}\dot{\vec{\tau}}(t) =$$
$$= [\tau_2\tau_3]^{-(\gamma-1)/2} \langle\!\langle 2\sqrt{\tau_1}, \tau_4-\tau_1, \tau_4-\tau_1, -2\sqrt{\tau_4} \rangle\!\rangle. \tag{3.5}$$

Согласно (1.18) это продолжимое по (3.5) семейство касательных векторов W раскладывается по базисным полям X_1, X_2, X_3, X_4 с коэффициентами, принимающими на кривой $\tau(t)$ вид

$$(W)_{\vec{\tau}(t)} = [t(1-t)]^{-\gamma}\{(2-2t)X_1 + (1-2t)(X_2+X_3) - 2tX_4\}_{\vec{\tau}(t)} \tag{3.6}$$

Это представление можно упростить по (1.19), взяв $c(t) = 1-2t = \tau_4 - \tau_1$:

$$\frac{d}{d\theta}\Big|_{\vec{\tau}(t)} = (W)_{\vec{\tau}(t)} = [t(1-t)]^{-\gamma}\{X_1 - X_4\}\Big|_{\vec{\tau}(t)}. \tag{3.7}$$

<u>Теорема 3.1</u>. Пусть ограничение векторного поля W, заданного по (3.5), на кривую $\vec{\tau}(t) = \vec{p}(t)\otimes\vec{p}(t)$ совпадает с дифференцированием $\frac{d}{d\theta}$ по геодезическому параметру $\theta = \overset{\gamma}{\theta}$ исходной инвариантной связности ${}^{\gamma}\nabla$, $\gamma\in\mathbb{R}$, на Cap_2. Тогда в этой же связности ${}^{\gamma}\nabla$ на $Cap_4 \supset Cap_2\otimes Cap_2$ перенос касательных векторов $(W)_{\vec{\tau}(t(\theta))}$ задается формулой

$${}^{\gamma}\nabla_W W\,\big|_{\vec{\tau}(t)} =$$
$$= -2\gamma(1-2t)\,[t(1-t)]^{-2\gamma}\{(1-t)X_1 + tX_4\}\Big|_{\vec{\tau}(t)} \tag{3.8}$$

Доказательство. По свойству символа ∇ из (3.7) следует

$$\nabla_W W = \nabla_W\,[t(1-t)]^{-\gamma}(X_1 - X_4) =$$
$$= \{W[t(1-t)]^{-\gamma}\}(X_1-X_4) + [t(1-t)]^{-\gamma}\nabla_W(X_1-X_4), \tag{3.9}$$

где для упрощения обозначений опущено, что рассмотрение ведется только на кривой $\vec{\tau}(t)$. Связь между дифференцированием $W = \frac{d}{d\theta}$

вдоль кривой и дифференцированием $\frac{d}{dt}$ навязывается их связью (2.6) в Cap_2 . Поэтому

$$\{W[t(1-t)]^{-\gamma}\}(X_1-X_4)=\left\{[t(1-t)]^{-\gamma+1}\frac{d}{dt}[t(1-t)]^{-\gamma}\right\}(X_1-X_4)=$$

$$=-\gamma\left[t(1-t)\right]^{-2\gamma}(1-2t)\ (X_1-X_4)\ . \tag{3.10}$$

Преобразуем теперь второе слагаемое:

$$[t(1-t)]^{-\gamma}\nabla_W(X_1-X_4)=[t(1-t)]^{-2\gamma}\nabla_{(X_1-X_4)}(X_1-X_4)=$$

$$=[t(1-t)]^{-2\gamma}\left\{\nabla_{X_1}X_1+\nabla_{X_4}X_4-\nabla_{X_1}X_4-\nabla_{X_4}X_1\right\}= \tag{3.11}$$

$$=\gamma\,[t(1-t)]^{-2\gamma}\{[1-2\tau_1(t)]X_1+[1-2\tau_4(t)]X_4-2[\tau_4(t)X_1+$$

$$+\tau_1(t)X_4\}=-\gamma\,[t(1-t)]^{-2\gamma}(1-2t)^2\,(X_1+X_4)\ .$$

Подставляя (3.10) и (3.11) в (3.9), получаем

$$\nabla_W W\ =$$

$$=-\gamma\,[t(1-t)]^{-\gamma}(1-2t)\left\{W-[t(1-t)]^{-\gamma}(1-2t)(X_1+X_4)\right\}. \tag{3.12}$$

Отсюда, приведя подобные члены, приходим к (3.8).

<u>Следствие</u>. Единственным кандидатом на наследственную инвариантную связность является связность ${}^0\nabla$ с $\gamma=0$.

В самом деле, векторные поля X_1 и X_4 линейно независимы, поэтому при $0<t<1$ вектор $(1-t)X_1+tX_4\neq 0$. Значит, при $\gamma\neq 0$ правая часть (3.8) обращается в нуль только при $t=1/2$. Это означает, что поле касательных векторов к кривой $\{\vec{\tau}(t(\theta))\}$ не параллельно вдоль $\{\vec{\tau}(t(\theta))\}$. При $\gamma=0$ правая часть (3.8) равна нулю тождественно.

Как утверждает теорема 21.1 из [1] , если $\{P_{\vec{s}}\}$ есть экспонентное семейство распределений с каноническим параметром $\vec{s}$, т.е. оно является вполне геодезической поверхностью в в инвариантной плоской линейной связности, отвечающей $\gamma = 0$, с геодезической параметризацией $\vec{s}$, то таково же семейство степеней $\{P_{\vec{s}}^{N}\}$ при любом $N \in \mathbb{N}$. Отсюда вытекает

<u>Теорема 3.2</u>. Логарифмическая плоская связность ${}^{0}\nabla$ является единственной инвариантной относительно категории марковских отображений линейной связностью, наследственной при переходе к степеням распределений.

Напомним, что только эта связность является наследственной также относительно переходов к условным распределениям, см. [1], теорема 12.3.

§ 4. <u>Параметризация геодезических</u>

Опишем в явном виде геодезическую линию в связности ${}^{\gamma}\nabla$ с $\gamma \neq 0$, проходящую по $\mathrm{Int}\ \mathrm{Cap}\,(\Omega_n, \mathcal{A}_n)$.

<u>Теорема 4.1</u>. Семейство векторов $\vec{p}\,(\theta)$,

$$p_j(\theta) = \left[\theta q_j^{\gamma} + (1-\theta)\, r_j^{\gamma}\right]^{1/\gamma} \left[H(\theta)\right]^{-1}, \quad j = 1, \ldots, n; \tag{4.1}$$

$$H(\theta) = \sum_{k=1}^{n} \left[\theta q_k^{\gamma} + (1-\theta)\, r_k^{\gamma}\right]^{1/\gamma}, \tag{4.2}$$

отвечает геодезической $\{P_\theta(\cdot)\}$, проходящей через $R = P_0$ и $Q = P_1$, $Q, R \in \mathrm{Int}\ \mathrm{Cap}\,(\Omega_n, \mathcal{A}_n)$. Отрезок $[\theta^-, \theta^+]$ изменения параметра θ, может быть полубесконечный, $\theta^- < 0 < 1 < \theta^+$, определяется пересечением условий

$$\theta q_j^{\gamma} + (1-\theta)\, r_j^{\gamma} \geqslant 0, \quad j = 1, \ldots, n. \tag{4.3}$$

Параметр θ не является, вообще говоря, каноническим. Геодезический параметр z определяется квадратурой

$$dz = C\left[H(\theta)\right]^{-2\gamma} d\theta\ ; \quad \forall\, C \in \mathbb{R} - \{0\}. \tag{4.4}$$

Доказательство. Продифференцировав (4.1), имеем

$$p_j'(\theta) = [\xi_j(\theta) - \eta(\theta)]\, p_j(\theta), \tag{4.5}$$

$$\xi_j(\theta) = \frac{1}{\gamma}\left(q_j^{\gamma} - r_j^{\gamma}\right)\left\{\theta q_j^{\gamma} + (1-\theta) r_j^{\gamma}\right\}^{-1}, \tag{4.6}$$

$$\eta(\theta) = H'(\theta)\,[H(\theta)]^{-1} = \frac{d}{d\theta}\ln H(\theta). \tag{4.7}$$

При дифференцировании тождественного равенства $\sum_{j=1}^{n} p_j(\theta) = 1$ получаем $\sum_{j=1}^{n} p_j'(\theta) = 0$, откуда, суммируя правую часть (4.5) по j, находим

$$\sum_{j=1}^{n} \xi_j(\theta)\, p_j(\theta) = \eta(\theta). \tag{4.8}$$

Отсюда для семейства векторов $(V)_{\vec{p}} = \frac{d}{d\theta}$ вдоль кривой (4.1) при $c = -\eta(\theta)$ по (1.19)

$$V = \sum_{j=1}^{n} \xi_j(\theta) X_j \quad . \tag{4.9}$$

Получим теперь две необходимые формулы. Дифференцируя (4.6) по параметру, устанавливаем

$$\frac{d\xi_j(\theta)}{d\theta} = -\gamma\,[\xi_j(\theta)]^2 \quad . \tag{4.10}$$

Запишем еще (2.3) и (2.4) единой формулой

$${}^{\gamma}\nabla_{X_k} X_j \Big|_{\vec{p}} = \gamma\,\delta_{jk} X_j \Big|_{\vec{p}} - \gamma\left[p_k X_j + p_j X_k\right]\Big|_{\vec{p}} \tag{4.11}$$

Перейдем к вычислению переноса:

$${}^{\gamma}\nabla_V V = {}^{\gamma}\nabla_V \left[\sum_{j=1}^{n} \xi_j(\theta) X_j\right] =$$

$$= \sum_{j=1}^{n} \left[V\,\xi_j(\theta)\right] X_j + \sum_{j=1}^{n} \xi_j(\theta)\cdot {}^{\gamma}\nabla_V X_j =$$

$$= \sum_{j=1}^{n} \left[\frac{d}{d\theta}\,\xi_j(\theta)\right] X_j + \sum_{j,k=1}^{n} \xi_j(\theta)\,\xi_k(\theta)\cdot {}^{\gamma}\nabla_{X_k} X_j =$$

$$= -\gamma \sum_{j=1}^{n} \left[\xi_j(\theta)\right]^2 X_j + \gamma \sum_{j=1}^{n} \left[\xi_j(\theta)\right]^2 X_j -$$

$$-2\gamma \sum_{j,k=1}^{n} \xi_k(\theta)\, p_k(\theta)\, \xi_j(\theta)\, X_j = -2\gamma\, \eta(\theta)\, V ,$$

где в ходе выкладок мы воспользовались (4.8) и (4.9). Таким образом, при бесконечно малом переносе поля касательных векторов вдоль параметризованной кривой (4.1) оно меняется, оставаясь пропорциональным себе,

$$ {}^{\gamma}\nabla_{V} V = -2\gamma\, \eta(\theta) V . \qquad (4.12)$$

Касательный вектор в параметризации (4.4) пропорционален исходному, а именно

$$W = \frac{d}{dz} = C\left[H(\theta)\right]^{2\gamma}\frac{d}{d\theta} = C\left[H(\theta)\right]^{2\gamma} V . \qquad (4.13)$$

Отсюда $\quad {}^{\gamma}\nabla_{W} W =$

$$= C^2\left[H(\theta)\right]^{4\gamma}\cdot \nabla_{V} V + 2\gamma\, C^2 \left[H(\theta)\right]^{4\gamma-1} H'(\theta) V =$$

$$= C^2 \left[H(\theta)\right]^{4\gamma} \left\{-2\gamma\, H'(\theta)\left[H(\theta)\right]^{-1} + 2\gamma\, H'(\theta)\left[H(\theta)\right]^{-1}\right\} V = 0 ,$$

что и требовалось показать.

В несколько иной форме сходные рассмотрения были проведены Амари [5] . При $\gamma = 0$ формулы (4.1) и (4.2), а также формула (4.6), содержат неопределенности, которые легко раскрыть, см. [5], и получить представление классического экспонентного семейства в канонической параметризации z , совпадающей в этом случае с θ , см. (4.4). Мы отсылаем за точными формулами к [1] , обращая внимание лишь на то, что $\theta^{\pm} = \pm\infty$ при $\gamma = 0$. Параметр θ будет каноническим также и при $\gamma = 1$, ибо тогда $H(\theta) = 1$

тождественно.

Еще один легко интегрируемый случай отвечает римановой геометрии, $\gamma = \frac{1}{2}$,

$$d\sigma = \frac{d\theta}{1 - 2\theta(1-\Sigma) + 2\theta^2(1-\Sigma)} , \qquad (4.14)$$

$$\Sigma = \sum_{k=1}^{n} \sqrt{\tau_k q_k} , \qquad (4.15)$$

ср. (2.11) и (2.12).

§ 5. Пучок геодезических

Самостоятельный интерес представляет описание геодезической $\{P_\theta\}$, проходящей через $R = P_0 \in \mathrm{Int\ Cap}\,(\Omega_n, \mathcal{A}_n)$ в данном направлении.

Теорема 5.1. Семейство векторов $\vec{p}\,(\vartheta)$,

$$p_j(\vartheta) = \tau_j\,(1 + \gamma\vartheta u_j)^{1/\gamma}\,[\mathcal{H}(\vartheta)]^{-1}, \quad j = 1,\dots,n; \qquad (5.1)$$

$$\mathcal{H}(\vartheta) = \sum_{k=1}^{n} \tau_k\,(1 + \gamma\vartheta u_k)^{1/\gamma} , \qquad (5.2)$$

отвечает геодезической $\{P_\vartheta(\cdot)\}$, проходящей через $R = P_0 \in \mathrm{Int\ Cap}\,(\Omega_n, \mathcal{A}_n)$, по направлению

$$\vec{u} = \frac{d \ln p_j(\vartheta)}{d\vartheta} ; \quad \sum_{k=1}^{n} u_k \tau_k = 0 . \qquad (5.3)$$

Отрезок $[\vartheta^-, \vartheta^+]$ изменения параметра ϑ, может быть полубесконечный, $\vartheta^- < 0 < \vartheta^+$, определяется пересечением условий

$$1 + \gamma\vartheta u_j \geq 0 , \qquad j = 1,\dots,n . \qquad (5.4)$$

Параметр ϑ не является, вообще говоря, каноническим, ср. теорему 4.1. Умножение вектора $\vec{u}$ на число приводит к той же геодезической, с пропорциональным изменением геодезического параметра. Последний по-прежнему связан с ϑ через нормирующий делитель $\mathcal{H}(\vartheta)$ аналогом (4.4):

$$dz = C\left[\mathcal{H}(\vartheta)\right]^{-2\gamma} d\vartheta . \tag{5.5}$$

<u>Лемма 5.2</u>. Пусть $\vartheta^- < \vartheta_0 < \vartheta^+$, $\vartheta_0 \neq 0$, $\vec{q} = \vec{\pi}(\vartheta_0)$. Тогда семейство $\{\vec{\pi}(\vartheta)\}$, см. (5.1) описывается семейством (4.1), т.е.

$$\vec{\pi}(\vartheta) = \vec{p}(\theta), \tag{5.6}$$

с дробно-линейной связью параметров θ и ϑ,

$$\theta = \vartheta\left[\mathcal{H}(\vartheta_0)\right]^{\gamma}\left\{\vartheta_0 - \vartheta + \vartheta\left[\mathcal{H}(\vartheta_0)\right]^{\gamma}\right\}^{-1}, \tag{5.7}$$

$$H(\theta) = \mathcal{H}(\vartheta)\,\vartheta_0^{1/\gamma}\left\{\vartheta_0 - \vartheta + \vartheta\left[\mathcal{H}(\vartheta_0)\right]^{\gamma}\right\}^{-1/\gamma}. \tag{5.8}$$

Доказательство. При $\vartheta = \vartheta_0$ из (5.1) имеем

$$q_j = r_j\left(1 + \gamma\vartheta_0 u_j\right)^{1/\gamma}\mathcal{H}_0^{-1}, \quad \forall j,$$

где $\mathcal{H}_0 = \mathcal{H}(\vartheta_0)$. Отсюда можно выразить u_j:

$$u_j = \frac{1}{\gamma\vartheta_0}\left[\frac{q_j^{\gamma}}{r_j^{\gamma}}\mathcal{H}_0^{\gamma} - 1\right], \quad \forall j. \tag{5.9}$$

Подставляя (5.9) в (5.1) после преобразований получаем при любом $\vartheta \in [\vartheta^-, \vartheta^+]$:

$$p_j(\vartheta) = \left\{r_j^{\gamma}\left(1 - \frac{\vartheta}{\vartheta_0}\right) + \frac{\vartheta}{\vartheta_0}\mathcal{H}_0^{\gamma} q_j^{\gamma}\right\}^{1/\gamma}\left[\mathcal{H}(\vartheta)\right]^{-1} =$$

$$= \left\{r_j^{\gamma}\frac{\vartheta_0 - \vartheta}{\mathcal{D}(\vartheta,\vartheta_0)} + q_j^{\gamma}\frac{\vartheta\,\mathcal{H}_0^{\gamma}}{\mathcal{D}(\vartheta,\vartheta_0)}\right\}^{1/\gamma}\cdot\left[\mathcal{D}(\vartheta,\vartheta_0)\right]^{1/\gamma}\left[\mathcal{H}(\vartheta)\right]^{-1}\vartheta_0^{-1/\gamma} =$$

$$= \left\{ r_j^{\gamma}(1-\theta) + q_j^{\gamma}\,\theta \right\}^{1/\gamma} \left[H(\theta)\right]^{-1} = p_j(\theta) ,$$

где для краткости обозначено

$$\mathcal{D}(\vartheta, \vartheta_0) = \vartheta_0 - \vartheta - \vartheta\,\mathcal{H}_0^{\gamma} , \qquad (5.10)$$

и коэффициент θ при q_j^{γ} связан с ϑ по (5.7), а нормирующий делитель $H(\theta)$ по единственности выражается через $\mathcal{H}(\vartheta)$ посредством (5.8). Этим лемма полностью доказана, откуда следует основное утверждение теоремы. Остается проверить формулу (5.5). Для этого продифференцируем (5.7), откуда

$$d\theta = \mathcal{H}_0^{\gamma}\,\vartheta_0 \left[\mathcal{D}(\vartheta, \vartheta_0)\right]^{-2} d\vartheta . \qquad (5.11)$$

Подставим теперь (5.8) и (5.11) в выражение (4.4) для геодезического параметра:

$$dz = C\left[H(\theta)\right]^{-2\gamma} d\theta =$$

$$= C\,\vartheta_0^{-2}\left[\mathcal{D}(\vartheta, \vartheta_0)\right]^{2}\left[\mathcal{H}(\vartheta)\right]^{-2\gamma}\mathcal{H}_0^{\gamma}\,\vartheta_0\left[\mathcal{D}(\vartheta, \vartheta_0)\right]^{-2} d\vartheta =$$

$$= \left\{C\,\vartheta_0^{-1}\left[\mathcal{H}(\vartheta_0)\right]^{\gamma}\right\}\left[\mathcal{H}(\vartheta)\right]^{-2\gamma} d\vartheta ,$$

что и требовалось показать.

Интересно было бы выяснить смысл величин $H(\theta)$ и $\mathcal{H}(\vartheta)$, как инвариантов, ср. [6].

§ 6. Вычисление кручения и кривизны

Так как и кручение, и кривизна являются тензорными полями, то достаточно вычислить их на базисных полях X_j. По определению кручения, см. [4], гл. I, § 8,

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X, Y] . \qquad (6.1)$$

Операторы инвариантной связности симметричны, см. (2.3), а базисные дифференцирования, согласно (1.21), коммутируют. Отсюда при $j \neq k$

$$T(X_j, X_k) = -\gamma\left(p_j X_k + p_k X_j\right) + \gamma\left(p_k X_j + p_j X_k\right) = 0.$$

При $j = k$ $T(X_j, X_j) = 0$ в силу кососимметричности определения (6.1). Отсюда вытекает

Теорема 6.1. Кручение любой инвариантной связности тождественно равно нулю,

$$^{\gamma}\nabla_X Y - {}^{\gamma}\nabla_Y X - XY + YX = 0. \qquad (6.2)$$

Прежде чем переходить к вычислению кривизны, построим явные выражения действия базисных полей X_k на базисные функции $p_j = p_j(\vec{p})$, где $\vec{p} = \langle\!\langle p_1, \dots, p_n \rangle\!\rangle$.

Лемма 6.2. Действие базисного поля X_k на базисную функцию p_j есть

$$X_k p_j = -p_k p_j, \quad j \neq k;$$

$$X_k p_k = p_k - p_k^2,$$

что можно записать единой формулой

$$X_k p_j = \delta_{jk} p_k - p_k p_j. \qquad (6.3)$$

Доказательство. В силу симметрии достаточно рассмотреть случай $k = 1$. Согласно (1.6) уравнение чевианы первого семейства имеет вид

$$p_1(t) = t,$$

$$p_j(t) = q_j(1-t), \quad j \geqslant 2,$$

где t - натуральный параметр вдоль чевианы, а $X = t(1-t)\frac{d}{dt}$, см. (2.6). Отсюда

$$X_1 p_1 = t(1-t) = p_1 - p_1^2,$$

$$X_1 p_j = q_j(1-t)t = -p_1 p_j, \quad j \geqslant 2,$$

что и требовалось доказать.

Рассмотрим теперь оператор кривизны

$$R(X,Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]} , \quad (6.4)$$

и его значения на базисных полях X_j :

$$R(X_i, X_j) X_k = R^{\ell}_{kij} X_\ell , \quad \forall\ k, i, j , \quad (6.5)$$

где по индексу ℓ подразумевается суммирование.

Теорема 6.2. Для инвариантной связности ${}^{\gamma}\nabla$ тензор кривизны R^{ℓ}_{kij} задается формулой

$$R(X_i, X_j) X_k = \gamma(\gamma-1) p_k \left[(p_j - \delta_{jk}) X_i - (p_i - \delta_{ik}) X_j \right]. \quad (6.6)$$

Следствие. Тензор ${}^{\gamma}R$ кривизны связности ${}^{\gamma}\nabla$ пропорционален тензору ${}^{1/2}R$ кривизны связности ${}^{1/2}\nabla$, отвечающей инвариантной римановой метрике,

$${}^{\gamma}R = 4\gamma(1-\gamma) \cdot {}^{1/2}R . \quad (6.7)$$

Доказательство. Так как по (1.21) базисные поля X_i и X_j коммутируют, нам достаточно вычислить ${}^{\gamma}\nabla_{X_i} {}^{\gamma}\nabla_{X_j} X_k$. Для унификации выкладки вместо выражений (2.3) и (2.4) воспользуемся единой формулой (4.11):

$${}^{\gamma}\nabla_{X_i} X_j = \gamma \delta_{ij} X_j - \gamma \left(p_i X_j + p_j X_i \right). \quad (6.8)$$

Последовательно вычисляем

$$\nabla_{X_j} X_k = \gamma \delta_{jk} X_k - \gamma p_j X_k - \gamma p_k X_j ;$$

$$\nabla_{X_i} \gamma \delta_{jk} X_k = \underline{\gamma^2 \delta_{ik} \delta_{jk} X_k} - \gamma^2 \delta_{jk} p_i X_k - \gamma^2 \delta_{jk} p_k X_i ,$$

$$- \nabla_{X_i} \gamma p_j X_k = \underline{-\gamma(\delta_{ij} p_i - p_i p_j) X_k} - \gamma^2 p_j \delta_{ik} X_k +$$

$$+\ \underline{\gamma^2 p_j p_i X_k} - \gamma^2 p_j p_k X_i \ ,$$

$$-\nabla_{X_i} \gamma p_k X_j = -\gamma\left(\delta_{ik} p_i - p_i p_k\right) X_j - \underline{\gamma^2 p_k \delta_{ij} X_j} +$$

$$+\ \underline{\gamma^2 p_k \left(p_i X_j + p_j X_i\right)} \ ,$$

где подчеркнутые прямой чертой слагаемые симметричны при перестановке i и j . Сложив последние три равенства, получаем

$$\nabla_{X_i} \nabla_{X_j} X_k = -\gamma^2 \delta_{jk} p_k X_i + \gamma^2 p_j p_k X_i -$$

$$-\gamma\left(\delta_{ik} p_i - p_i p_k\right) X_j + \ \ldots \qquad ,$$

где точками обозначены пропущенные симметричные слагаемые. При этом учтено, что два слагаемых, подчеркнутые волнистой чертой, имеют симметричную сумму. Окончательно,

$$\nabla_{X_i} \nabla_{X_j} X_k = -\gamma p_k \left[\gamma(\delta_{jk} - p_j) X_i + (\delta_{ik} - p_i) X_j\right] + \ldots \ , \tag{6.9}$$

где использовано, что $\delta_{ik} p_i = \delta_{ik} p_k$. Вычитая из (6.9) аналогичное выражение для $\nabla_{X_j} \nabla_{X_i} X_k$, получающееся перестановкой индексов i и j , приходим к (6.6). Формула (6.7) является очевидным следствием (6.6). Доказательство завершено.

ЛИТЕРАТУРА

1. Ченцов Н.Н. Статистические решающие правила и оптимальные выводы. М.: Наука, 1972.
2. Ченцов Н.Н. Общая теория статистического вывода. Дисс. д.ф.-м.н. М.: ИПМ им. М.В.Келдыша АН СССР, 1968.
3. Cencov N.N. Algebraic Foundation of Mathematical Statistics. - Math. Operationsforsch. Statist., Ser. Statistics. 1972, Vol. 9, N 2, 267-276.
4. Helgason S. Differential Geometry and Symmetric Spaces. Academic Press, 1962.
5. Amari Shun-ichi. Differential-Geometrical Methods in Statistics. Lecture Notes in Statistics, Vol. 28. Springer, 1985.
6. Морозова Е.А., Ченцов Н.Н. Марковские инварианты пар вероятностных законов. Препринт // М. Ин-т прикл. матем. им. М.В. Келдыша АН СССР, 1986, № 121.

Е.А. Морозова, Н.Н. Ченцов " Инвариантные линейные связности на совокупностях распределений вероятностей."
Редактор М.И. Граев. Корректор Ю.Б. Котов.

Подписано в печать 5.09.88г. № Т-14066. Заказ № 363.
Формат бумаги 60Х90 1/16. Тираж 185 экз.
Объём 1,1 уч.-изд.л. Цена 15 коп.

055 (02)2

Отпечатано на ротапринтах в Институте прикладной математики АН СССР

Москва, Миусская пл. 4.

Цена 15 коп.