

Information Geometry and Its Applications

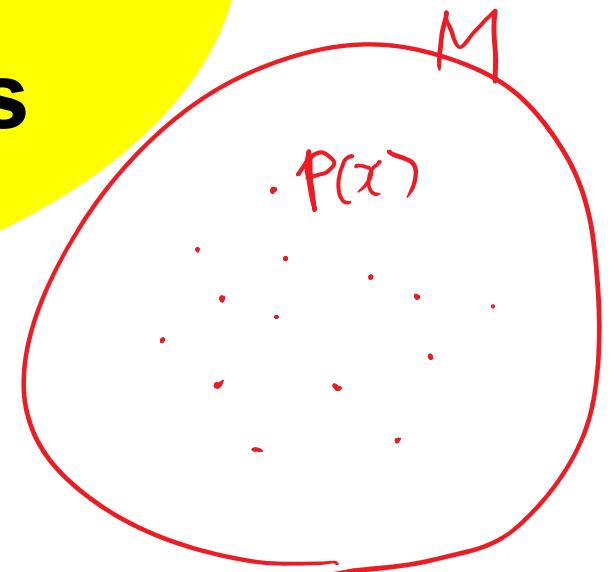
Shun-ichi Amari RIKEN Brain Science Institute

1. Divergence Function and Dually Flat Riemannian Structure
2. Invariant Geometry on Manifold of Probability Distributions
3. Geometry and Statistical Inference
 - semi-parametrics
4. Applications to Machine Learning and Signal Processing

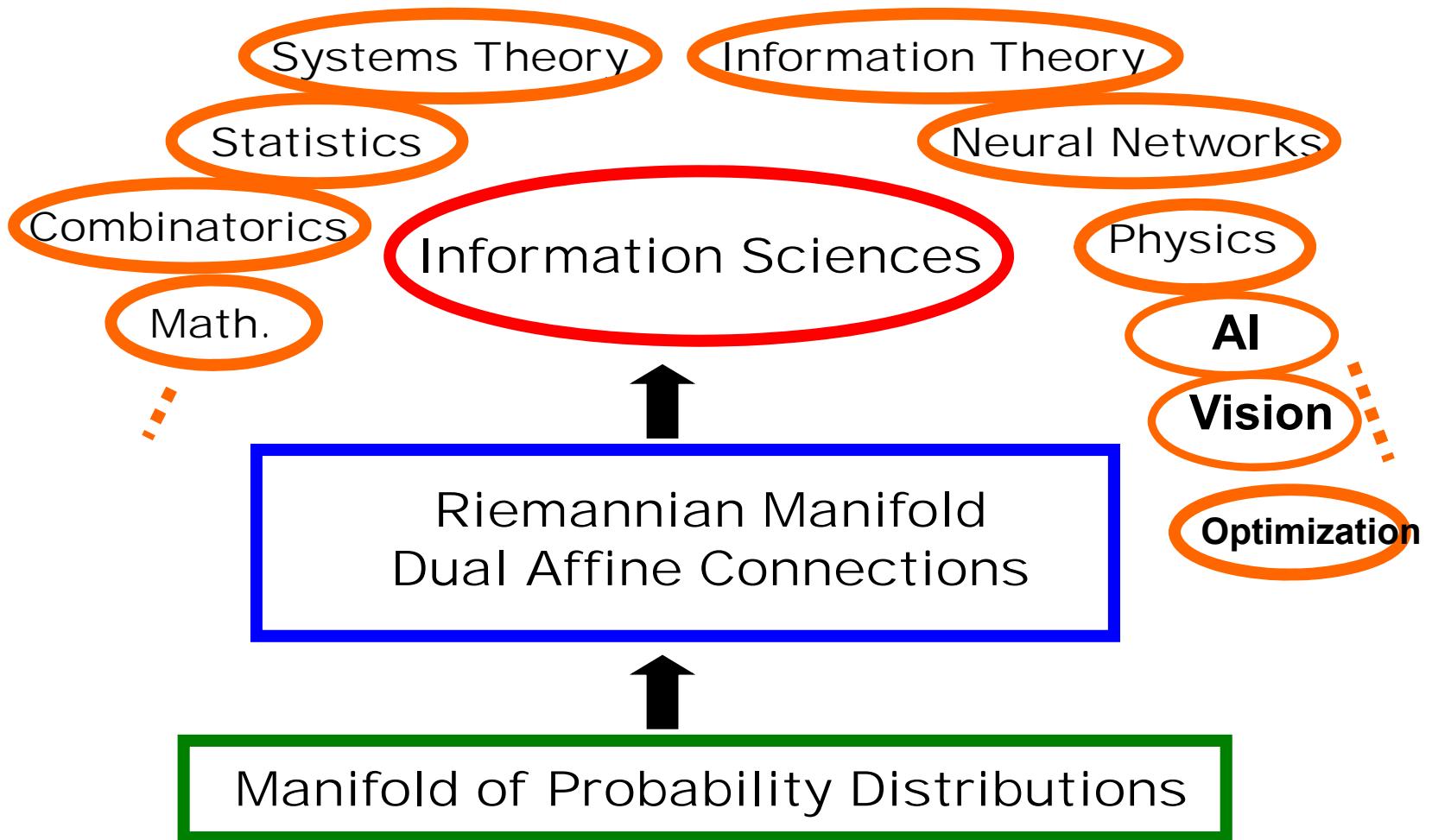
Information Geometry

-- Manifolds of
Probability Distributions

$$M = \{p(\mathbf{x})\}$$



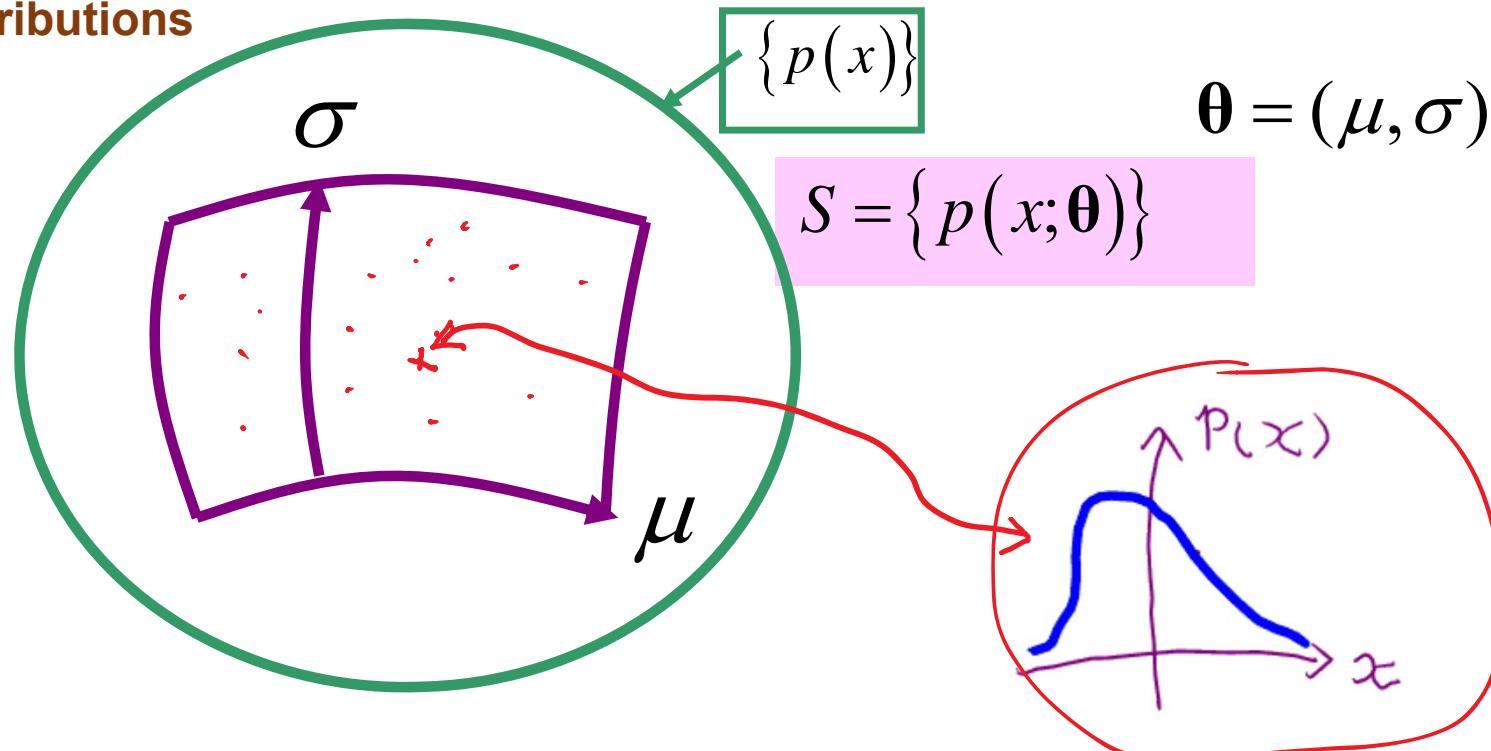
Information Geometry



Information Geometry ?

$$S = \{ p(x; \mu, \sigma) \} \quad p(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

Gaussian distributions

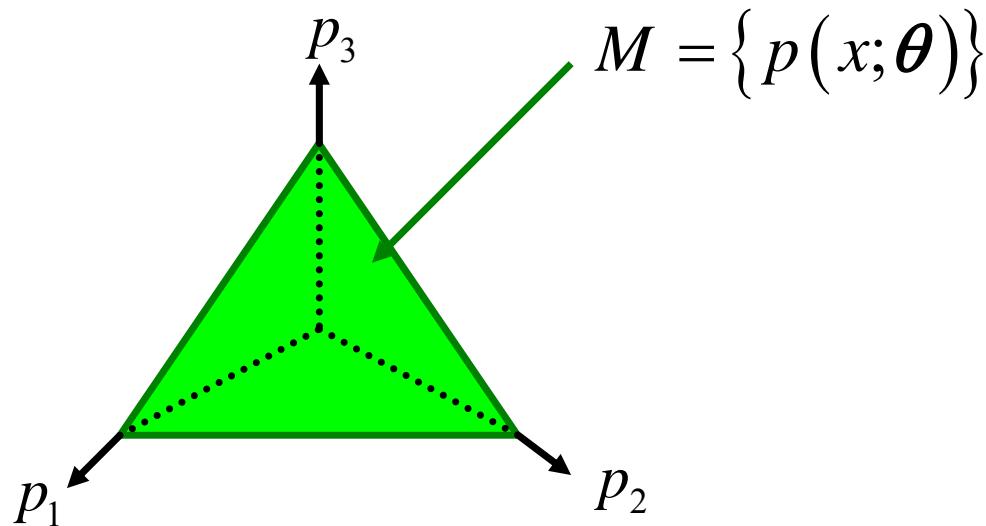


Manifold of Probability Distributions

$$x = 1, 2, 3 \quad S_n = \{p(x)\}$$

n = 3

$$p = (p_1, p_2, p_3), \quad \boxed{p_1 + p_2 + p_3 = 1}$$



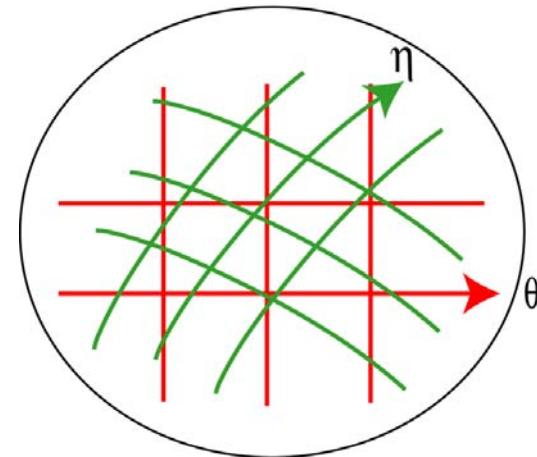
Manifold and Coordinate System

coordinate transformation

$$\theta = (\theta_1, \dots, \theta_n)$$

$$\eta = (\eta_1, \dots, \eta_n)$$

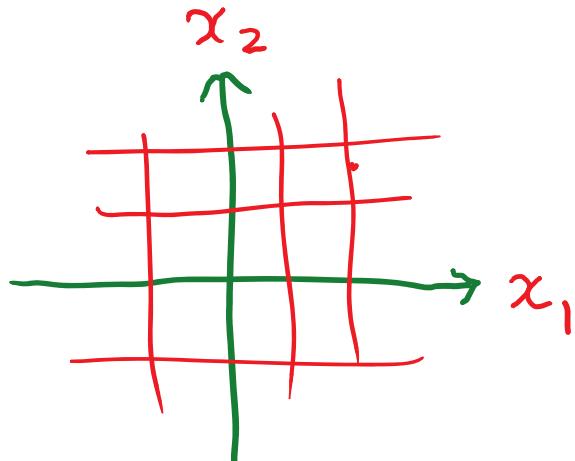
$$\eta = \eta(\theta) \leftrightarrow \theta = \theta(\eta)$$



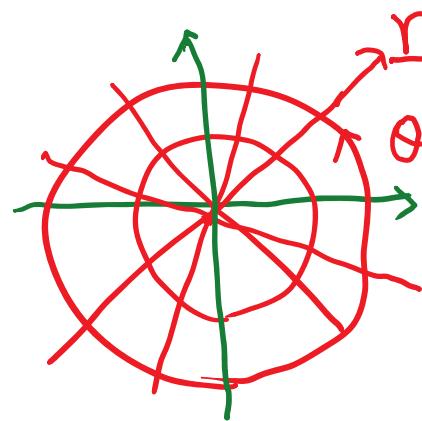
one-to-one
differentiable

Examples of Coordinate systems

Euclidean space



$$\mathbf{x} = (x_1, x_2)$$



$$P = (r, \theta)$$

$$\begin{cases} x_1 = r \cos \theta \\ x_2 = r \sin \theta \end{cases}$$

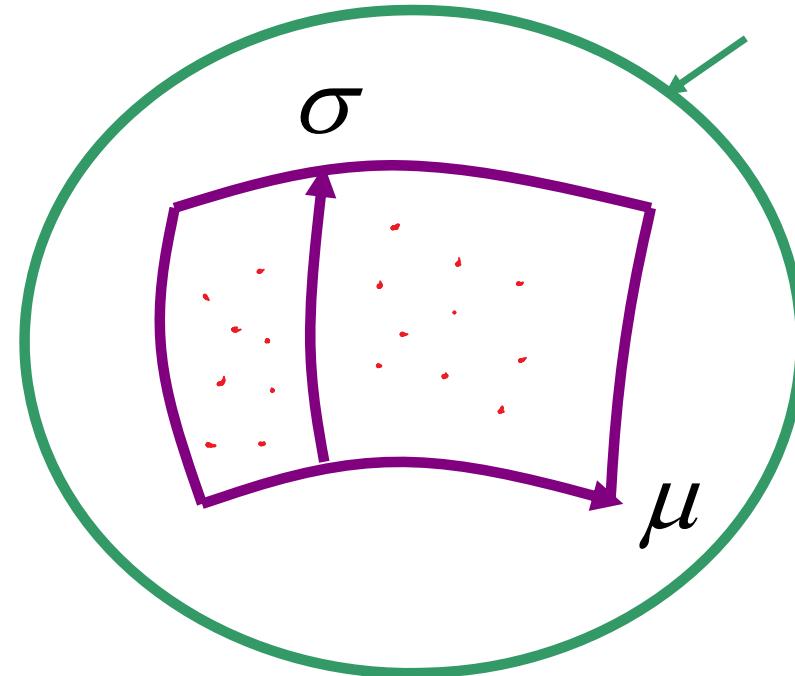
$$\begin{cases} r = \sqrt{x_1^2 + x_2^2} \\ \theta = \tan^{-1} \frac{x_2}{x_1} \end{cases}$$

Gaussian distributions

$$\mathfrak{X} = (\mu, \sigma^2),$$

$$\Theta = \left(-\frac{1}{2\sigma^2}, \frac{\mu^2}{\sigma^2} \right),$$

$$\eta = (\mu, \mu^2 + \sigma^2)$$

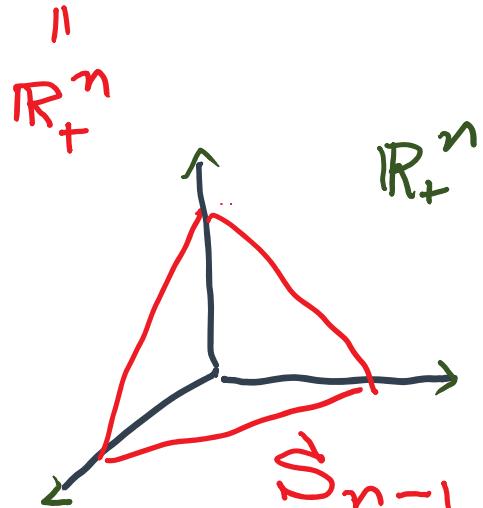


$$S = \{ p(x; \mu, \sigma) \} \quad p(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}$$

Discrete Distributions $S_n = \{P(x)\}, x = 0, 1, \dots, n$

Positive measures

$$M_n : m(x) = \sum m_i \delta_i(x)$$



$$P(x) = \sum_{i=0}^n p_i \delta_i(x)$$

$$= \sum_{i=1}^n p_i \delta_i(x) - p_0$$

$$\eta = (p_1, \dots, p_n)$$

$$\theta = \left(\log \frac{p_1}{p_0}, \dots, \log \frac{p_n}{p_0} \right)$$

.

Divergence: $D[z : y]$

$$D[z : y] \geq 0$$

$$D[z : y] = 0, \quad \text{iff } z = y$$

Not necessarily symmetric

$$D[z : y] \neq D[y : z]$$

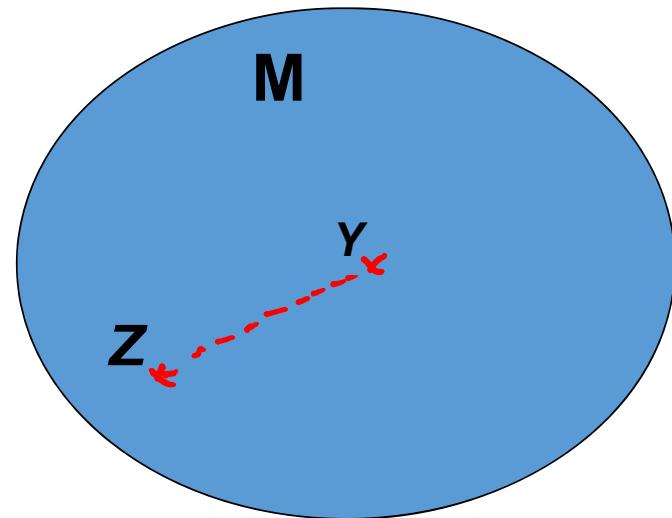
$$D[z : z + dz] = \frac{1}{2} \sum g_{ij} dz_i dz_j$$

positive-definite $G = (g_{ij})$

Taylor expansion

$$D(z : z + dz) = \frac{1}{2} \sum g_{ij} dz_i dz_j + \frac{1}{6} \sum k_{ijk} dz_i dz_j dz_k$$

+ ...



Various Divergences

Euclidean

$$\mathcal{D}[x : y] = \frac{1}{2} \sum (x_i - y_i)^2 \quad : \text{symmetric}$$

f-divergence

$$\mathcal{D}[p(x) : q(x)] = \int p(x) f\left\{\frac{q(x)}{p(x)}\right\} dx$$

KL-divergence

f : convex function ,
 $f(1) = 0$

(α - β)-divergence

Kullback-Leibler Divergence

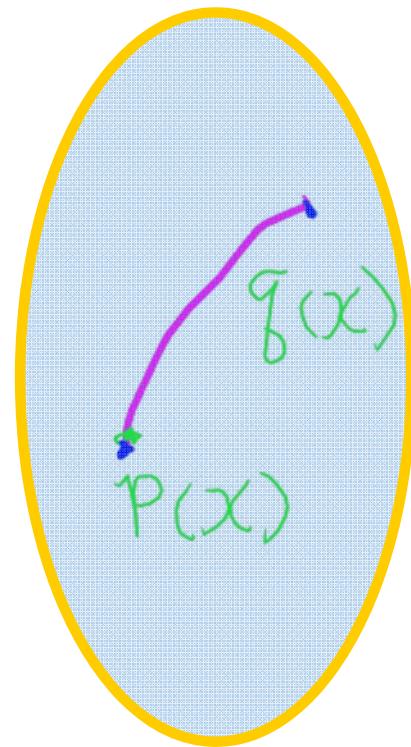
quasi-distance

$$f(u) = -\log(u)$$

$$D[p(x) : q(x)] = \sum_x p(x) \log \frac{p(x)}{q(x)}$$

$$D[p(x) : q(x)] \geq 0 \quad = 0 \text{ iff } p(x) = q(x)$$

$$D[p : q] \neq D[q : p]$$



(α, β) -divergence

$P, Q \in S_n$

$$D_{\alpha, \beta}[P : Q] = \sum_i \left\{ \frac{\alpha}{\alpha + \beta} p_i^{\alpha + \beta} + \frac{\beta}{\alpha + \beta} q_i^{\alpha + \beta} - p_i^\alpha q_i^\beta \right\}$$

$\beta = -\alpha$: α -divergence

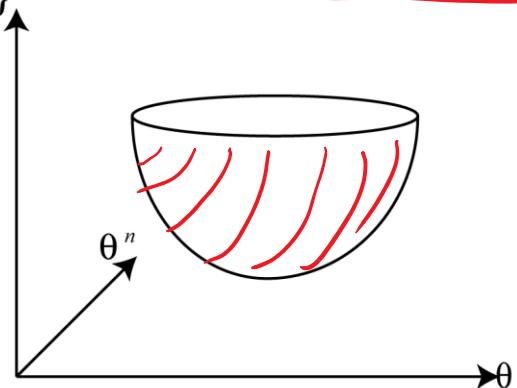
$\alpha = 1$: β -divergence

\

Manifold with Convex Function

S : coordinates $\theta = (\theta^1, \theta^2, \dots, \theta^n)$

$\psi(\theta)$: convex function



$$\psi(\theta) = \frac{1}{2} \sum (\theta^i)^2$$

**negative entropy
energy**

$$\varphi(p) = \int p(x) \log p(x) dx$$

**mathematical programming, control systems
physics, engineering, vision, economics**

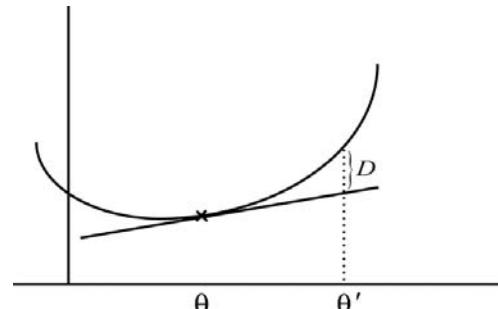
Riemannian metric and flatness (affine structure)

$$\{S, \psi(\theta), \theta\}$$

convex: $\hat{\theta} = A\theta + c$

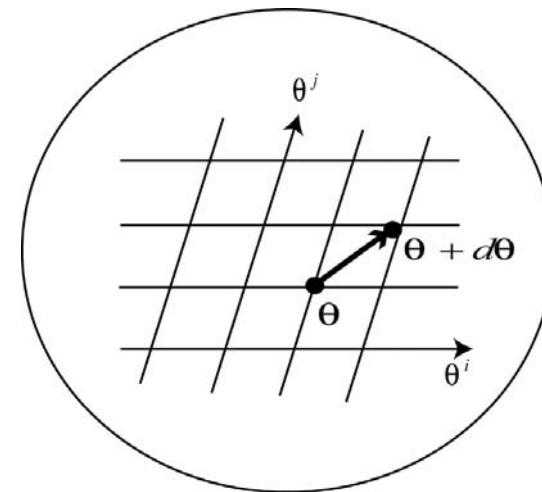
Bregman divergence

$$D(\theta, \theta') = \psi(\theta') - (\theta' - \theta) \cdot \text{grad } \psi(\theta)$$



$$D(\theta, \theta + d\theta) = \frac{1}{2} \sum g_{ij}(\theta) d\theta^i d\theta^j$$

$$g_{ij} = \partial_i \partial_j \psi(\theta), \quad \partial_i = \frac{\partial}{\partial \theta^i}$$



Flatness (affine) θ : geodesic (not Levi-Civita)

$\theta(t) = t\alpha + b$

Legendre Transformation

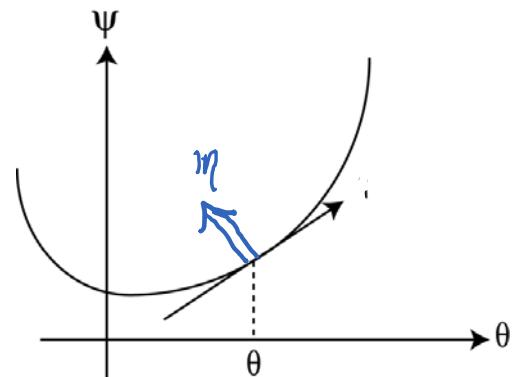
$$\eta_i = \partial_i \psi(\theta), \quad \partial_i = \frac{\partial}{\partial \theta^i}$$

$\psi(\theta) \leftrightarrow \theta \leftrightarrow \eta \leftrightarrow \varphi(\eta)$ one-to-one

$$\varphi(\eta) + \psi(\theta) - \theta_i \eta^i = 0$$

$$\theta^i = \partial^i \varphi(\eta), \quad \partial^i = \frac{\partial}{\partial \eta_i}$$

dual coordinates (θ, η)



$$\varphi(\eta) = \max_{\theta} \{ \theta^i \eta^i - \psi(\theta) \}$$

$$D(\theta, \theta') = \psi(\theta) + \varphi(\eta') - \theta \cdot \eta' \quad : \text{proof easy}$$

Two affine coordinate systems (θ, η)

θ : geodesic (e-geodesic)

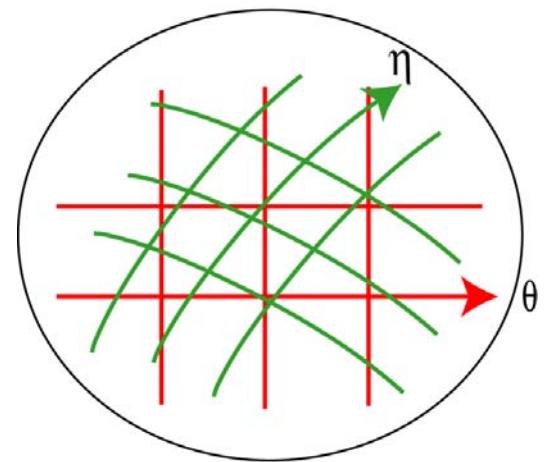
η : dual geodesic (m-geodesic)

“dually orthogonal”

$$\langle \partial_i, \partial^j \rangle = \delta_i^j$$

$$\partial_i = \frac{\partial}{\partial \theta^i}, \quad \partial^i = \frac{\partial}{\partial \eta_i}$$

$$X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X^* Z \rangle$$



Bi-orthogonality

$$d\theta = \sum d\theta^i e_i$$

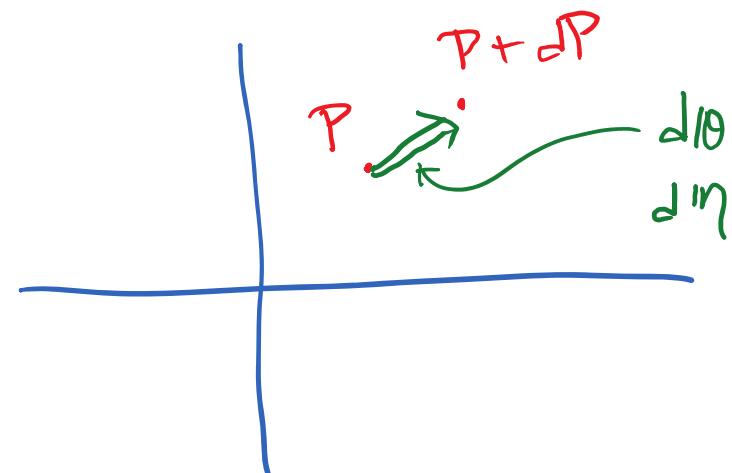
$$d\eta = \sum d\eta_i e^i$$

$$\eta_i = \partial_i \varphi(\theta)$$

$$d\eta_i = \sum \partial_i \partial_j \varphi(\theta) d\theta^j = \sum g_{ij} d\theta^j$$

$$d\theta^i = \sum g^{ij} d\eta_j$$

$$(g_{ij})^{-1} = \delta^i \delta^j \varphi(\eta)$$



$$ds^2 = \langle d\theta, d\theta \rangle$$

$$= \sum \langle e_i, e_j \rangle d\theta^i d\theta^j$$

$$e_i = \sum g_{ij} e^j$$

$$\langle e_i, e^j \rangle = \delta_i^j$$

Dually flat manifold

θ -coordinates \leftrightarrow η -coordinates

potential functions $\psi(\theta), \varphi(\eta)$

$$g_{ij}(\theta) = \frac{\partial^2}{\partial \theta_i \partial \theta_j} \psi(\theta) \dots g^{ij} = \frac{\partial^2}{\partial \eta_i \partial \eta_j} \varphi(\theta)$$

$$\psi(\theta) + \varphi(\eta) - \sum \theta_i \eta_i = 0$$

exponential family: $p(x, \theta) = \exp \left\{ \sum \theta_i x_i - \psi(\theta) \right\}$

ψ : cumulant generating function

φ : negative entropy

canonical divergence $D(P: P') = \psi(\theta) + \varphi(\eta') - \sum \theta_i \eta'_i$

Exponential Family

$$p(x, \theta) = \exp\{\theta \cdot x - \psi(\theta)\} \quad \psi(\theta): \text{convex function, free-energy}$$

Gaussian:

$$\begin{cases} \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x \\ x^2 \end{pmatrix}, & \theta = \begin{pmatrix} \theta' \\ \theta^2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2\sigma^2} \\ \frac{\mu^2}{\sigma^2} \end{pmatrix} \\ \theta \cdot \boldsymbol{x} = -\frac{(x-\mu)^2}{2\sigma^2} + c \end{cases}$$

Negative entropy

$$\varphi(\eta) = \int p(\boldsymbol{x}, \theta) \log p(\boldsymbol{x}, \theta)$$

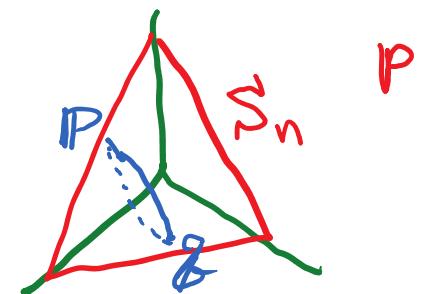
natural parameter : θ

expectation parameter : $\eta = E[\boldsymbol{x}]$

x : discrete $X = \{0, 1, \dots, n\}$

$S_n = \{p(x) \mid x \in X\}$: exponential family

$$p(x) = \sum_{i=0}^n p_i \delta_i(x) = \exp \left[\sum_{i=1}^n \theta^i x_i - \psi(\theta) \right] = \exp \left\{ \theta \cdot x - \psi(\theta) \right\}$$



$$\log M(t) = (1-t) \log P + t \log Q$$

$$\theta^i = \log(p_i / p_0); \quad x_i = \delta_i(x); \quad \psi(\theta) = -\log p_0$$

$$\eta_i = E[x_i] = p_i \quad \varphi(\eta) = \sum p_i \log p_i \quad M(t) = (1-t)P + tQ$$

Two geodesics

$$\theta(t) = t a_1 + b$$

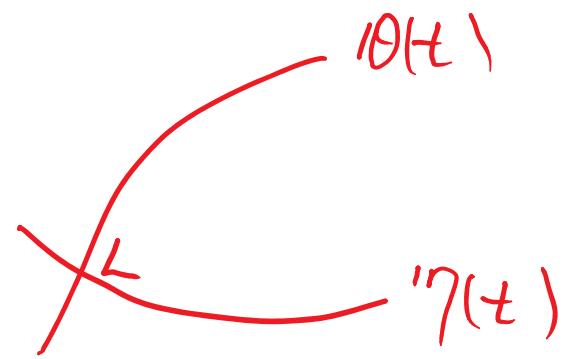
$$\eta(t) = t a' + b$$

Tangent directions

$$\dot{\theta}(t) = \sum \dot{\theta}^i e_i$$

$$\dot{\eta}(t) = \sum \dot{\eta}_j e^j$$

$$\langle \dot{\theta}, \dot{\eta} \rangle = \sum \dot{\theta}^i \dot{\eta}_i = 0$$



orthogonal

:

Function space of probability distributions: topology $\{p(x)\}$

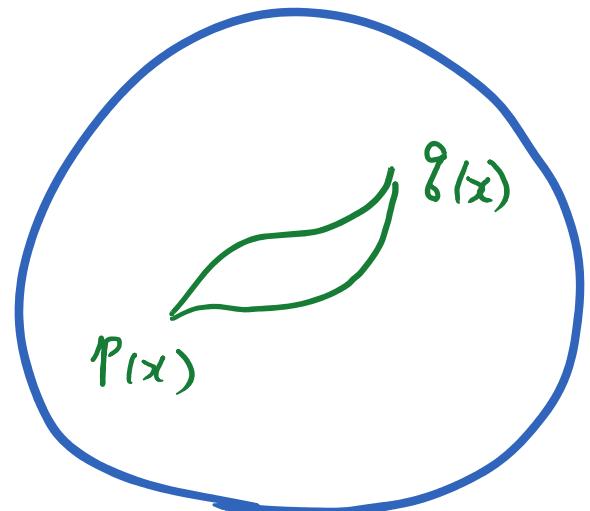
$$\mathcal{F} = \{ p(x) > 0, \quad \int p(x) dx = 1 \}$$

m-geodesic

$$r_m(x, t) = (1-t)p(x) + t\delta(x)$$

e-geodesic

$$\log r_e(x, t) = (1-t)\log p(x) + t\log \delta(x) + \zeta(t)$$



Exponential Family

$$\mathcal{F} = \{ p(x) \}$$

$$p(x) = \exp \left\{ \int \theta(s) \delta(s-x) ds - \psi(\theta) \right\}$$

$$\theta(s) = \log p(s), \quad \psi(\theta) = \log \left\{ \exp \{ \theta(s) \} ds \right\}$$

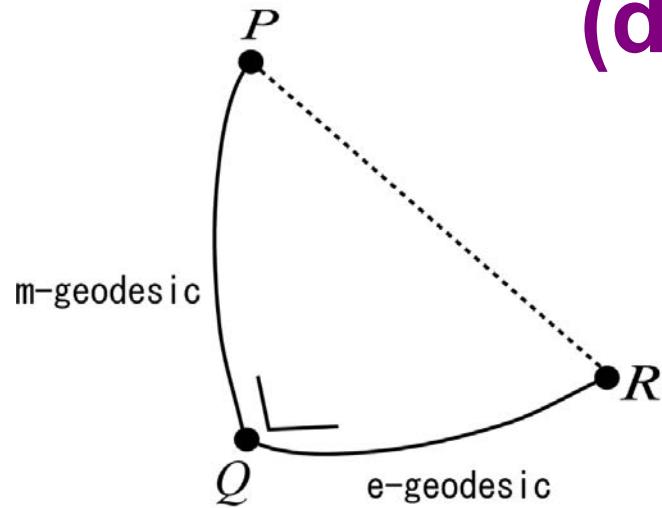
$$\eta(s) = E[\delta(s-x)] = p(s) = \nabla \psi(\theta)$$

$$D_{KL}[p : g] = \int p(x) \log \frac{p(x)}{g(x)} dx$$

$$g(s,t) = \nabla \nabla \psi = p(s) \delta(s-t)$$

Pythagorean Theorem

(dually flat manifold)



$$D[P:Q] + D[Q:R] = D[P:R]$$

proof

Euclidean space: self-dual $\theta = \eta$

$$\psi(\theta) = \frac{1}{2} \sum (\theta_i)^2$$

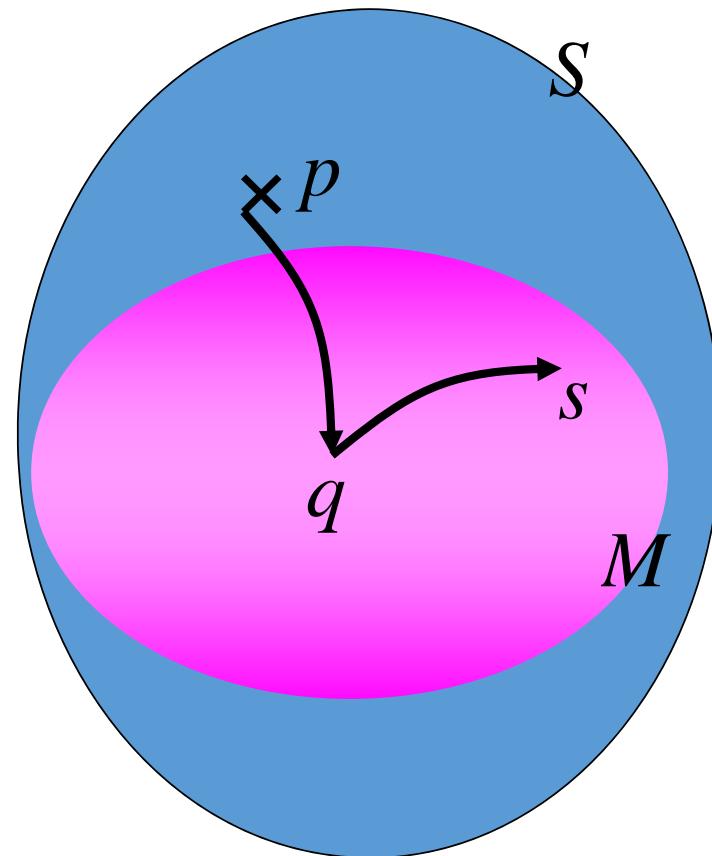
Projection Theorem

$$q = \arg \min_{s \in M} D[p : s]$$

m-geodesic

$$q = \arg \min_{s \in M} D[s : p]$$

e-geodesic



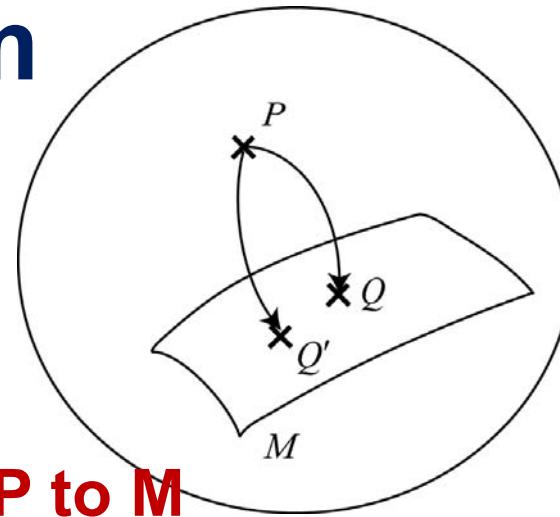
Projection Theorem

$$\min_{Q \in M} D[P : Q]$$

Q = m-geodesic projection of P to M

$$\min_{Q \in M} D[Q : P]$$

Q' = e-geodesic projection of P to M



Convex function – Bregman divergence
– Dually flat Riemannian divergence

$$\psi(\theta) \Rightarrow D_\psi[\theta : \theta'] \Rightarrow \{\theta, \eta\}, \quad g = \nabla D\psi$$

Dually flat R-manifold – convex function – canonical divergence
KL-divergence

$$\begin{aligned} \{\theta, \eta\} &\Rightarrow \psi(\theta) \Rightarrow D_\psi[\theta : \theta'] \\ g = \frac{\partial \eta}{\partial \theta} &. \end{aligned}$$
$$D_{KL}[p(x) : g(x)]$$

Exponential family – Bregman divergence

Banerjee et al

$$p(x, \theta) = \exp\{\theta \cdot x - \psi(\theta)\} \Rightarrow D_\psi(\theta : \theta')$$

$$D_\psi[\theta : \theta'] = \psi(\theta) + \varphi(x') - \theta \cdot x' , \quad \eta' = x'$$

$$\Rightarrow p(x, \theta) = \exp\{-D_\psi[\theta : \theta(x)] + \psi(x)\}$$

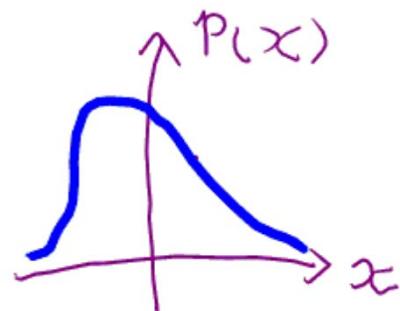
Invariance

$$S = \{ p(x, \theta) \}$$

Invariant under different representation of x

$$y = y(x), \quad \bar{p}(y, \theta)$$

$$\int |p(x, \theta_1) - p(x, \theta_2)|^2 dx \\ \neq \int |\bar{p}(y, \theta_1) - \bar{p}(y, \theta_2)|^2 dy$$



Invariant divergence

(manifold of probability
distributions; $S = \{p(x, \xi)\}$)

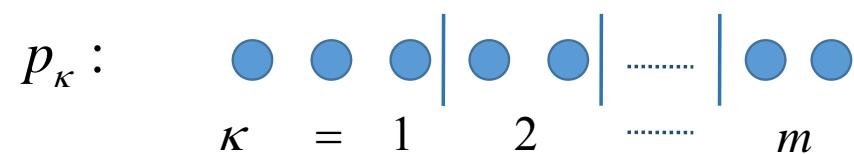
Chentsov
Amari -Nagaoka

$y = k(x)$: sufficient statistics

$$D[p_x(x) : q_x(x)] = D[p_y(y) : q_y(y)]$$

Invariance

--- characterization of f -divergence



$$p^A = (p_\kappa)$$

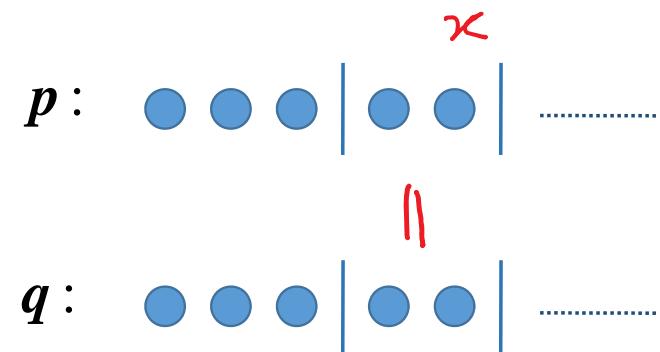
$$p_\kappa = \sum_{i \in A_\kappa} p_i$$

$\mathbb{P} \rightarrow \mathbb{P}^A$

$$D[p : q] \geq D[p^A : q^A]$$

$$D[p : q] = D[p^A : q^A]$$

$$\Leftrightarrow p_i = c_\kappa q_i ; i \in A_\kappa$$



Invariance $\Rightarrow f$ -divergence

Csiszar f -divergence

Ali-Silvey
Morimoto

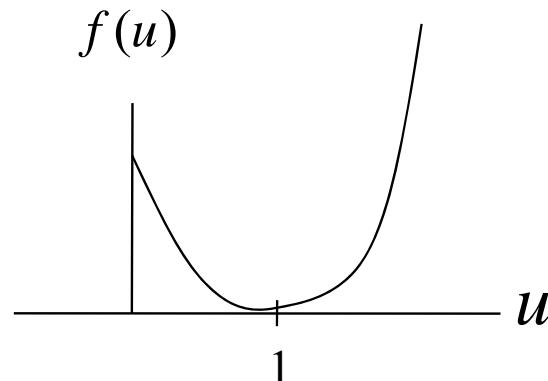
$$D_f [p : q] = \sum p_i f\left(\frac{q_i}{p_i}\right),$$

$$f(u) : \text{convex}, \quad f(1) = 0,$$

$$D_{cf} [p : q] = c D_f [p : q]$$

$$\tilde{f}(u) = f(u) - c(u-1)$$

$$f(1) = f'(1) = 0 ; f''(1) = 1$$



Theorem

An invariant separable divergence belongs to the class of f-divergence.

Separable divergence:

$$D[\mathbf{p} : \mathbf{q}] = \sum k(p_i, q_i)$$

$$k(p_i, q_i) = p_i f\left(\frac{q_i}{p_i}\right)$$

divergence

$S = \{p\}$: space of probability distributions

invariance

invariant divergence

F-divergence
Fisher inf metric
Alpha connection

dually flat space

Flat divergence

convex functions
Bregman

KL-divergence

$$D[p : q] = \int p(x) \log\left\{\frac{p(x)}{q(x)}\right\} dx$$

α -Divergence: why?

flat & invariant in $\tilde{S}_{n+1} = \mathcal{M}_{n+1}$

$$f_\alpha(u) = \frac{4}{1-\alpha^2} \left\{ 1 - u^{\frac{1+\alpha}{2}} \right\} - \frac{2}{1-\alpha} (1-u), \quad \alpha \neq 1$$

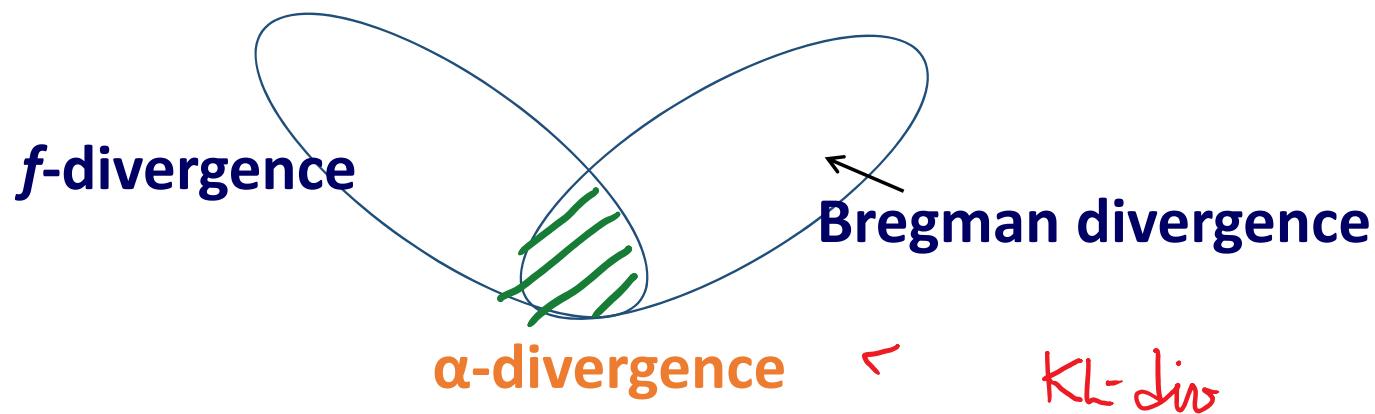
KL-divergence

$$f(u) = u \log u - (u - 1)$$

$$D[\tilde{p} : \tilde{q}] = \sum \left\{ \tilde{p}_i \log \frac{\tilde{p}_i}{\tilde{q}_i} + \tilde{p}_i - \tilde{q}_i \right\}$$

Space of positive measures : vectors, matrices, arrays

$$\tilde{S} = \{\tilde{p}\}, \quad \tilde{p}_i > 0 : (\sum \tilde{p}_i = 1 \text{ } nn \text{ holds})$$



f divergence of \tilde{S}

$$D_f[\tilde{p} : \tilde{q}] = \sum \tilde{p}_i f\left(\frac{\tilde{q}_i}{\tilde{p}_i}\right) \geq 0$$

$$D_f[\tilde{p} : \tilde{q}] = 0 \Leftrightarrow \tilde{p} = \tilde{q}$$

not invariant under $\tilde{f}(u) = f(u) - c(u-1)$

α divergence

$$D_{\alpha}[\tilde{p} : \tilde{q}] = \sum \left\{ \frac{1-\alpha}{2} \tilde{p}_i + \frac{1+\alpha}{2} \tilde{q}_i - \tilde{p}_i^{\frac{1-\alpha}{2}} \tilde{q}_i^{\frac{1+\alpha}{2}} \right\}$$

if θ_i // η_i

KL-divergence

$$D[\tilde{p} : \tilde{q}] = \sum \left\{ \tilde{p}_i \log \frac{\tilde{p}_i}{\tilde{q}_i} + \tilde{p}_i - \tilde{q}_i \right\}$$

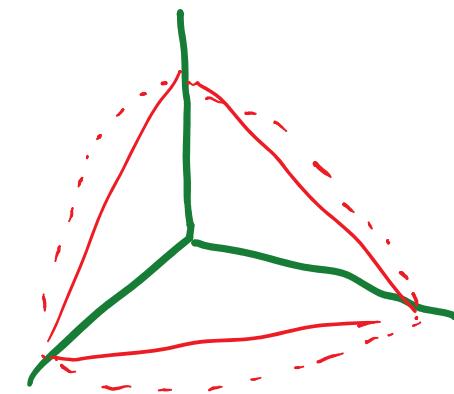
α

$M = \tilde{S}$: dually flat

S : not dually flat (except $\alpha = \pm 1$)

$$\sum p_i = 1$$

$$\sum r_i^{\frac{2}{1-\alpha}} = 1$$



Metric and Connections Induced by Divergence

Riemannian metric

(Eguchi)

$$g_{ij}(z) = \partial_i \partial_j D[z : y]_{|y=z} : D[z : y] = \frac{1}{2} g_{ij}(z)(z_i - y_i)(z_j - y_j)$$

affine connections

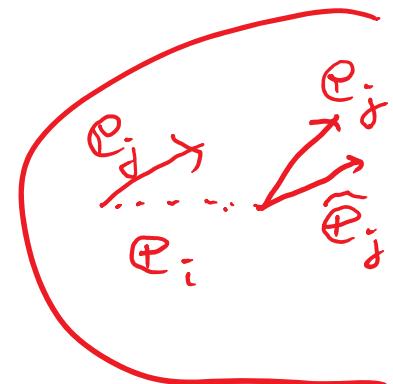
$\{\nabla, \nabla^*\}$

$$\nabla_{\hat{e}_i} \hat{e}_j = \sum \Gamma_{ij}^k \hat{e}_k$$

$$\Gamma_{ijk}(z) = -\partial_i \partial_j \partial_k' D[z : y]_{|y=z}$$

$$\partial_i = \frac{\partial}{\partial z_i}, \quad \partial_i' = \frac{\partial}{\partial y_i}$$

$$\Gamma_{ijk}^*(z) = -\partial_i' \partial_j' \partial_k D[z : y]_{|y=z}$$



Invariant geometrical structure alpha-geometry (derived from invariant divergence)

$$S = \{ p(x, \xi) \}$$

$$g_{ij}(\xi) = E[\partial_i l \partial_j l] \quad \text{Fisher information}$$

$$T_{ijk}(\xi) = E[\partial_i l \partial_j l \partial_k l] \quad l = \log p(x, \xi); \quad \partial_i = \frac{\partial}{\partial \xi^i}$$

α -connection

$$\Gamma_{ijk}^\alpha = \{i, j; k\} - \alpha T_{ijk} \quad \text{Levi-civita:}$$

$$\nabla^\alpha \leftrightarrow \nabla^{-\alpha} \quad \text{: dually coupled}$$

$$X \langle Y, Z \rangle = \langle \nabla_x^\alpha Y, Z \rangle + \langle Y, \nabla_x^{-\alpha} Z \rangle$$

Duality: $X < Y, Z > = < \nabla_X Y, Z > + < Y, \nabla^*_X Z >$

$$\partial_k g_{ij} = \Gamma_{kij} + \Gamma_{kji}^*$$

$$\Gamma_{ijk} = \Gamma_{ijk}^* - T_{ijk}$$

$$\boxed{\{M, g, T\}}$$

$$g_{ij}, T_{ijk}$$

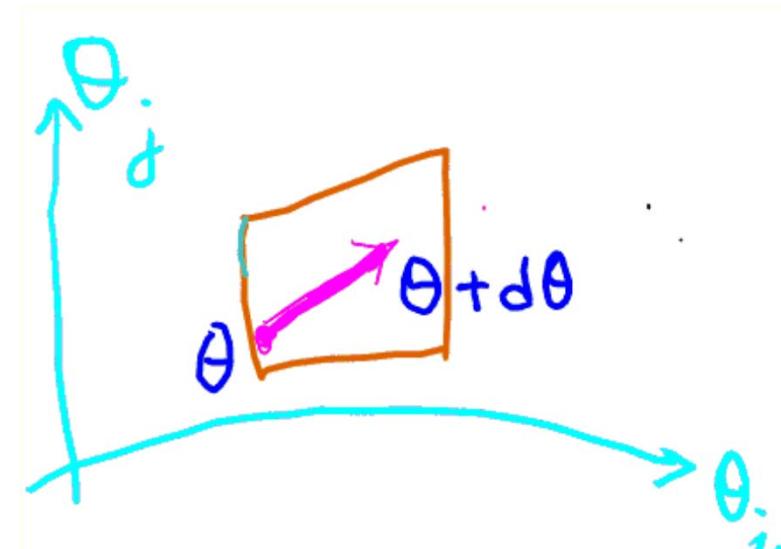
Riemannian Structure

$$\begin{aligned}ds^2 &= \sum g_{ij}(\theta) d\theta^i d\theta^j \\&= d\theta^T G(\theta) d\theta\end{aligned}$$

$$G(\theta) = (g_{ij})$$

Euclidean $G = E$

Fisher information



Affine Connection

covariant derivative

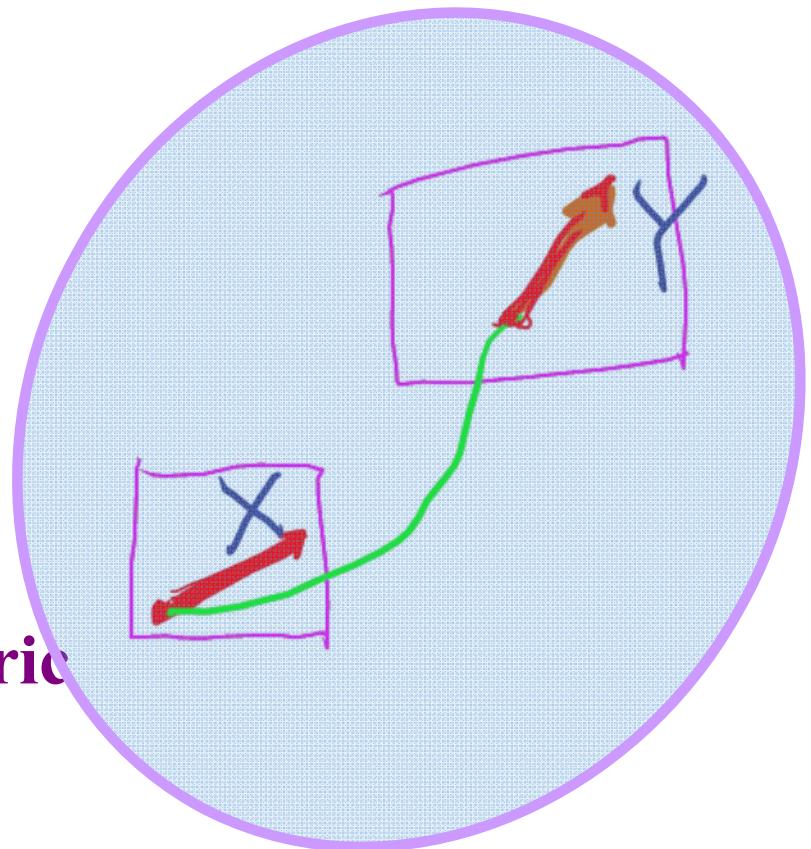
$$\nabla_X Y, \quad \Pi_c X = Y$$

geodesic $\nabla_{\dot{X}} \dot{X} = 0, \quad X=X(t)$

$$s = \int \sqrt{\sum g_{ij}(\theta) d\theta^i d\theta^j}$$

minimal distance \Rightarrow non-metric

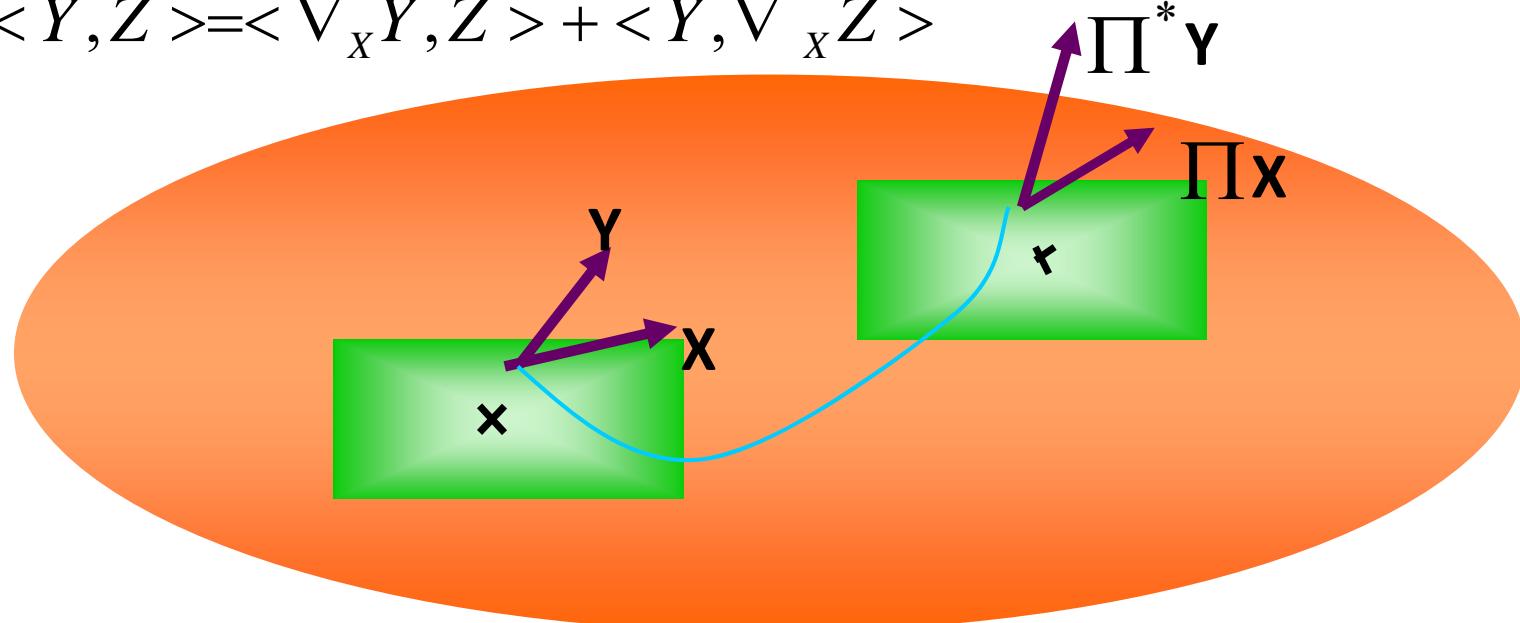
straight line



Duality

$$\langle X, Y \rangle = \langle \Pi X, \Pi^* Y \rangle \quad \langle X, Y \rangle = \sum g_{ij} X^i Y^j$$

$$X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla^*_X Z \rangle$$



Riemannian geometry: $\Pi = \Pi^*$

Dual Affine Connections

$$(\nabla, \nabla^*)$$

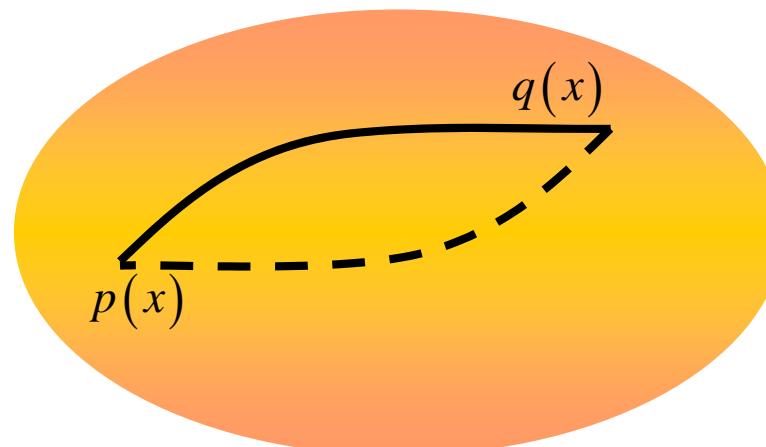
$$(\Pi, \Pi^*)$$

e-geodesic

$$\log r(x, t) = t \log p(x) + (1-t) \log q(x) + c(t)$$

m-geodesic

$$r(x, t) = tp(x) + (1-t)q(x)$$



Mathematical structure of $S = \{p(x, \xi)\}$

$$g_{ij}(\xi) = E[\partial_i l \partial_j l]$$
$$T_{ijk}(\xi) = E[\partial_i l \partial_j l \partial_k l]$$
$$\{\mathbf{M}, \mathbf{g}, \mathbf{T}\}$$

$$l = \log p(x, \xi); \quad \partial_i = \frac{\partial}{\partial \xi^i}$$

α -connection

$$\Gamma_{ijk}^\alpha = \{i, j; k\} - \alpha T_{ijk}$$

$\nabla^\alpha \leftrightarrow \nabla^{-\alpha}$: dually coupled

$$X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X^* Z \rangle$$

α -geometry

$$\tilde{\Gamma}^{(\alpha)} = \Gamma^{(0)} - \frac{\alpha}{2} T$$

not flat

$$(\alpha \neq \pm 1)$$

α -geodesic
 α -projection

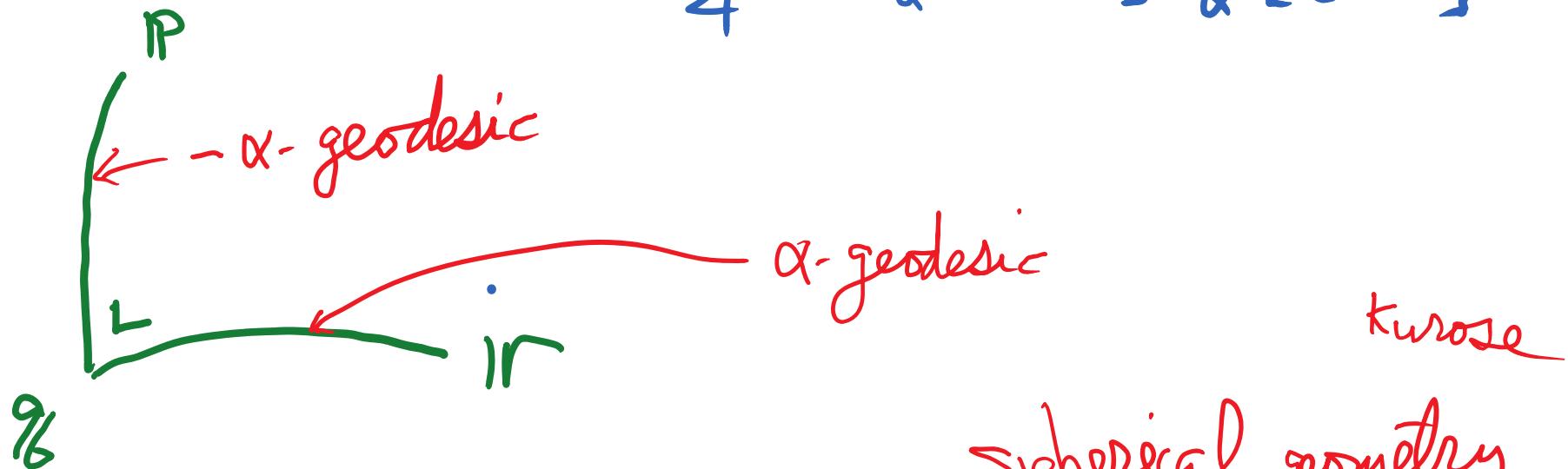
extended Pythagorean theorem

α -mean, α -family of prob. distributions

Tsallis q -entropy

$$D_\alpha [P : \text{ir}] = D_\alpha [P : g] + D_\alpha [g : \text{ir}]$$

$$+ \frac{1 - \alpha^2}{4} D_\alpha [P : g] D_\alpha [g : \text{ir}]$$



spherical geometry
(constant curvature)

Dual Foliations

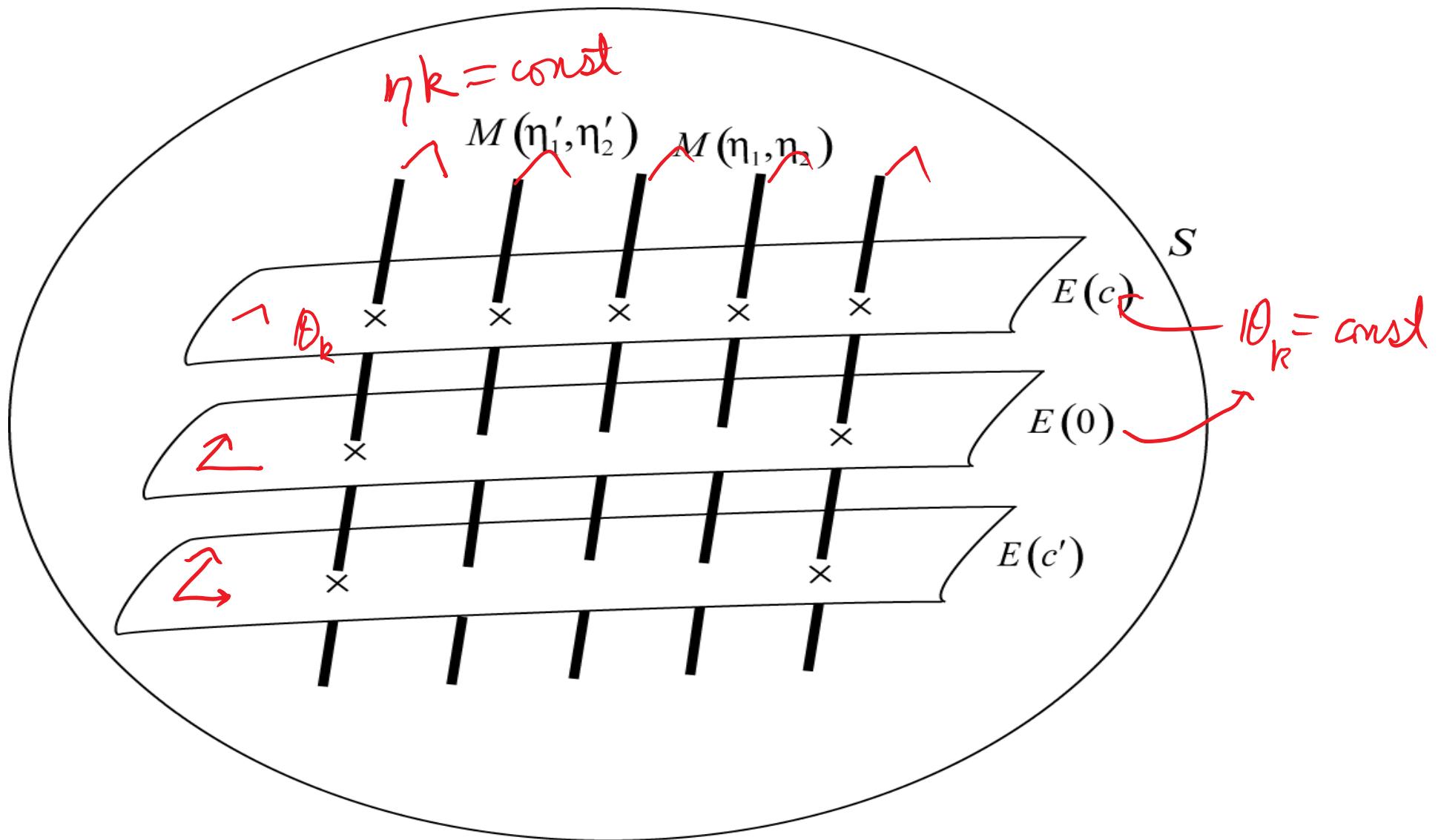
$$\theta = (\theta_1, \dots, \theta_k; \theta_{k+1}, \dots, \theta_n)$$

$$\eta = (\eta_1, \dots, \eta_k; \eta_{k+1}, \dots, \eta_n)$$

k-cut

$$\xi = (\theta_k : \eta^k)$$

mixed coordinates



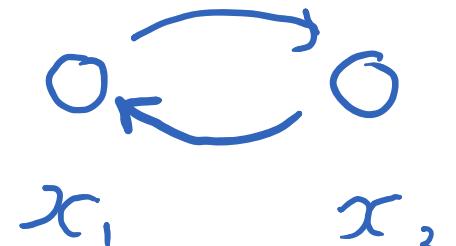
Two neurons: $\{p_{00}, p_{01}, p_{10}, p_{11}\}$

x_1

0011000101101

x_2

0100100110100



x_3

0101101001010

firing rates: $r_1, r_2; r_{12}$
correlation—covariance?

Correlations of Neural Firing

$$\{p(x_1, x_2)\}$$

$$\{p_{00}, p_{10}, p_{01}, p_{11}\}$$

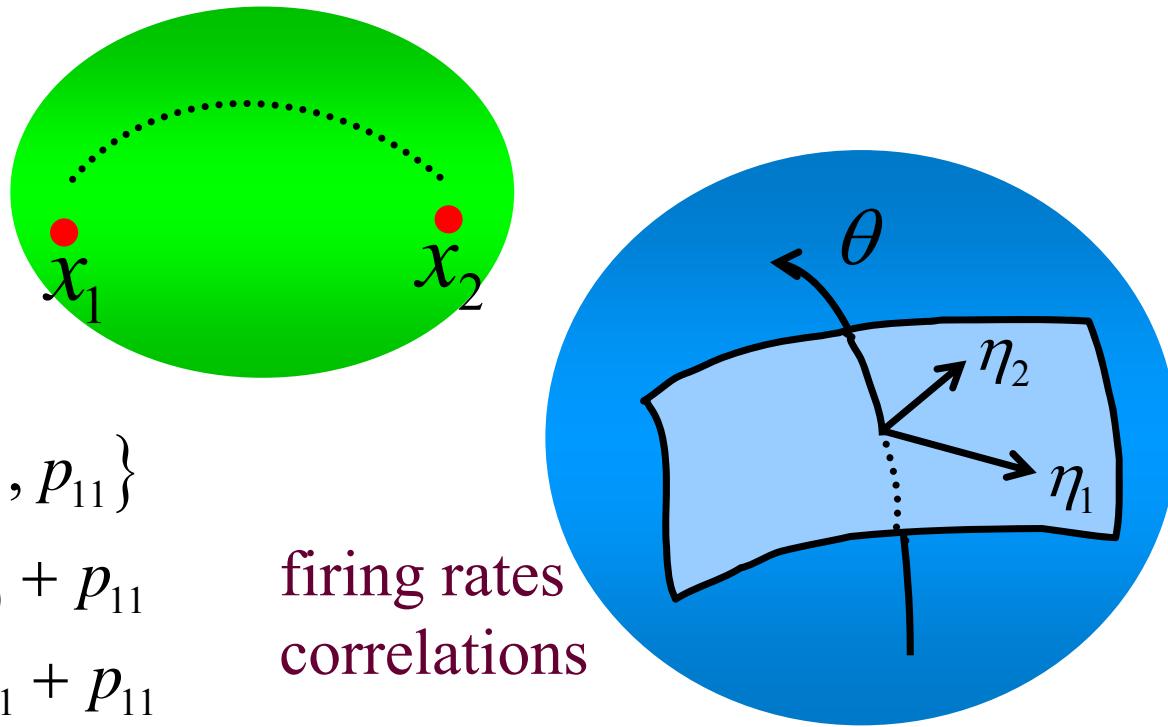
$$r_1 = p_{1\cdot} = p_{10} + p_{11}$$

$$r_2 = p_{\cdot 1} = p_{01} + p_{11}$$

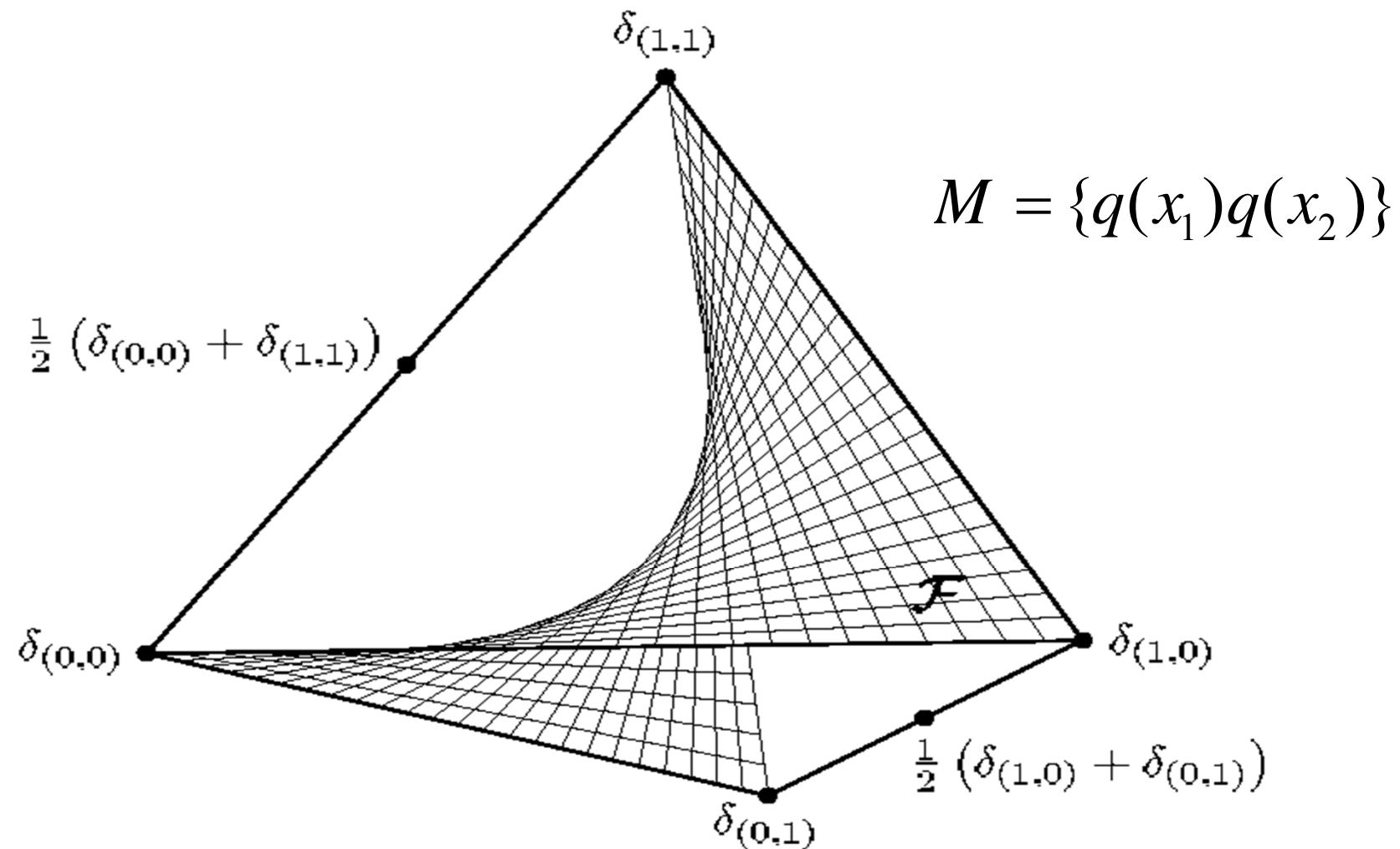
$$\theta = \log \frac{p_{11}p_{00}}{p_{10}p_{01}}$$

firing rates
correlations

$\{(r_1, r_2), \theta\}$
orthogonal coordinates



Independent Distributions



two neuron case

$$r_1, r_2, r_{12}; \theta_1, \theta_2, \theta_{12}$$

$$\theta_{12} = \log \frac{p_{00}p_{11}}{p_{01}p_{10}} = \log \frac{r_{12}(1 + r_{12} - r_1 - r_2)}{(r_1 - r_{12})(r_2 - r_{12})}$$

$$r_{12} = f(r_1, r_2, \theta)$$

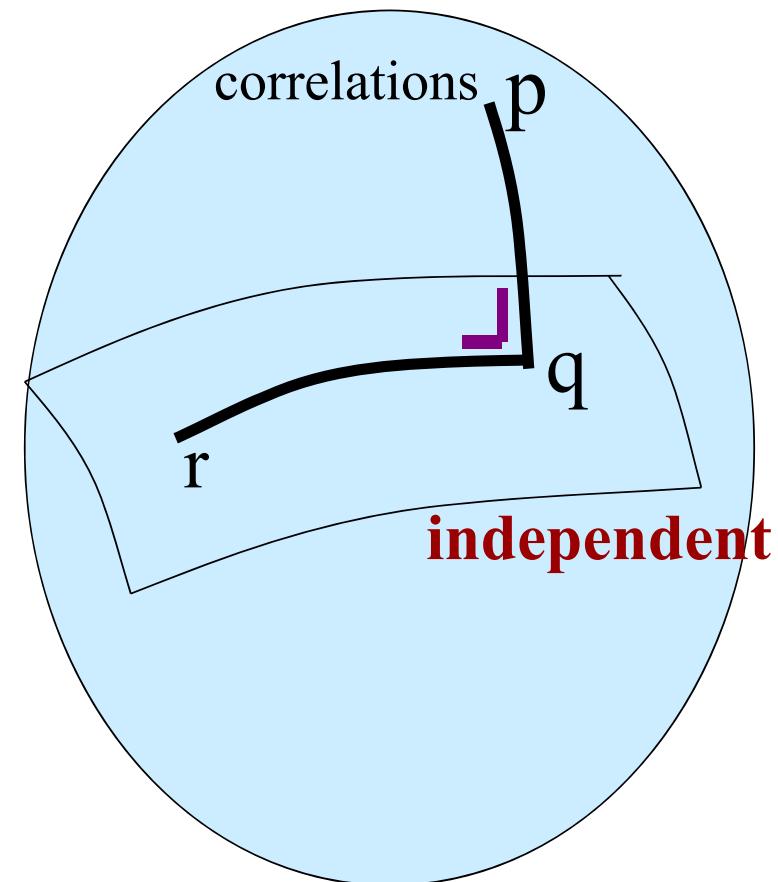
$$r_{12}(t) = f(r_1(t), r_2(t), \theta)$$

Decomposition of KL-divergence

$$D[p:r] = D[p:q] + D[q:r]$$

p,q: same marginals η_1, η_2
r,q: same correlations θ

$$D[p:r] = \sum_x p(x) \log \frac{p(x)}{q(x)}$$



pairwise correlations

covariance: $c_{ij} = r_{ij} - r_i r_j$ not orthogonal

independent distributions

$$r_{ij} = r_i r_j, \quad r_{ijk} = r_i r_j r_k, \dots$$

How to generate correlated spikes?
(Niebur, Neural Computation [2007])

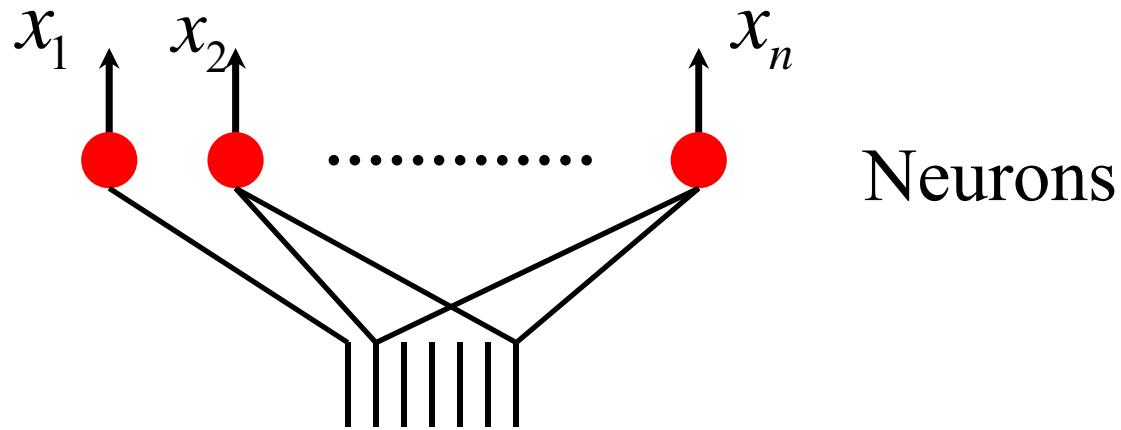
higher-order correlations

Orthogonal higher-order correlations

$$\theta = \left(\theta_i, \theta_{ij}; \dots, \theta_{1\dots n} \right)$$

$$r = \left(r_i, r_{ij}; \dots, r_{1\dots n} \right)$$

Population and Synfire



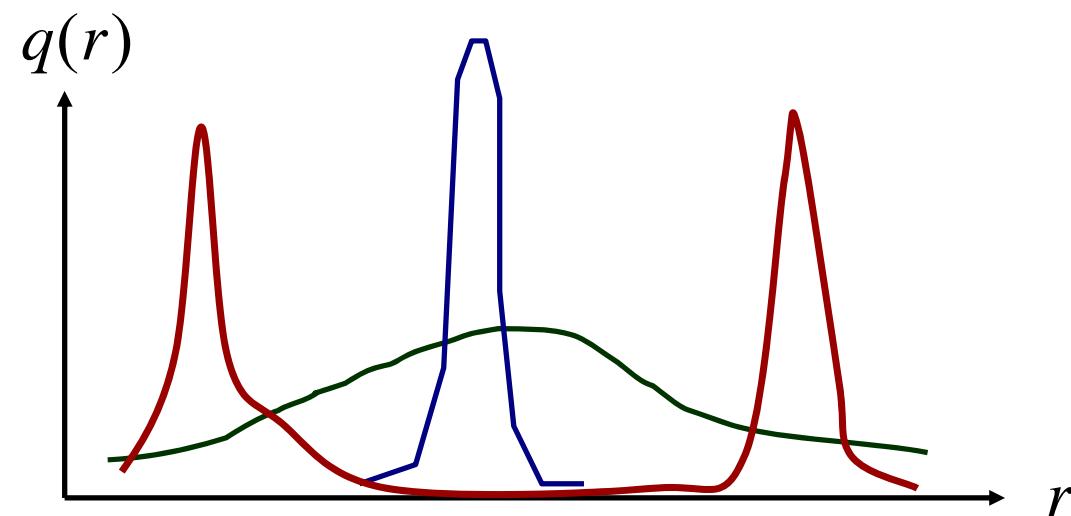
$$x_i = \text{1}\left(u_i\right)$$

$$u_i = \text{Gaussian} \quad E[u_i u_j] = \alpha$$

Synfiring

$$p(x) = p(x_1, \dots, x_n)$$

$$r = \frac{1}{n} \sum x_i \qquad q(r)$$



Input-output Analysis

Gross product consumption

Relations among industries

(K. Tsuda and R. Morioka)

$$(A_{ij}) = \begin{bmatrix} & 1 & \dots & n \\ \vdots & A_{11} & A_{12} & \dots & A_{1n} \\ & \ddots & \ddots & \ddots & \vdots \\ n & A_{n1} & & & A_{nn} \end{bmatrix}$$

n industries

product

$$A_{i\cdot} = \sum_j A_{ij}$$

consumption

$$A_{\cdot j} = \sum_i A_{ij}$$

$$S_{ij} = \log \frac{A_{ij}}{A_{i\cdot} A_{\cdot j}}$$

T_{θ_k}

Mathematical Problems

M: submanifold of S_n ?

Hong van Le

n: large

yes

{M, g} {M, g, T} dually flat: J. Armstrong

no

Affine differential geometry

Hessian manifold

Almost complex structure