Morse Lemma. (Normal forms) differential topology enables one to analyze the topology of a manifold by studying differentiate functions on that manifold.

Diffeomorphisms 微分同胚　光滑：可限可微

bijective 双射 mapping : $F: U \to V$, F^{-1} denoting the inverse mapping.

定义：

F is C^k - diffeomorphism ($k \geq 1$) if $F \in C^k(U, V)$ and $F^{-1} \in C^k(V, U)$

↓ Continuous differentiable. C^1：一阶连续可微　C^0：连续

导数存在且连续

Optimality Conditions 最优性条件

优化问题中：如果 x^* 是局部最优解

它必须满足什么条件？

本质："一个解是最优的，需要满足某些方程/条件"

思想：calculus of variations 变分法

the method of Lagrange multipliers：寻找 local maxima and minima

(Multivariable Calculus)

极值问题的必要条件

拉格朗日乘子法

-morphism：形式，结构

isomorphism 同构

homeomorphism 同胚

diffeomorphism 微分同胚

例子：maximize $f(x,y) = xy$

Subject to $3x + y = 720$　　$g(x,y) = 3x + y - 720 = 0$

$$F(x,y,\lambda) = f(x,y) + \lambda g(x,y) = xy + \lambda(3x + y - 720)$$

$F_x(x,y,\lambda) = 0 \Rightarrow y + 3\lambda = 0$

$F_y(x,y,\lambda) = 0 \Rightarrow x + \lambda = 0$　$\Rightarrow \lambda = -120$　$x = 120$, $y = 360$.

$F_\lambda(x,y,\lambda) = 0 \Rightarrow 3x + y - 720 = 0$

The solutions are critical points of F.

PCA 优化最大方差：$\max u^T S u$

Subject to $\|u\| = 1$

$$F(u,\lambda) = u^T S u + \lambda(1 - u^T u)$$

$F_u = u^T S + \lambda u^T = 0$　　$u^T S = -\lambda u^T \Rightarrow S^T u = -\lambda u$

$F_\lambda = u^T u - 1 = 0$

对称 $\Rightarrow Su = -\lambda u$

eigenvalue.

Lagrange Multipliers：把"有约束问题"变成：

$$L(x,\lambda) \equiv f(x) + \lambda g(x)$$

Known as the Lagrangian function. or just Lagrangian.

把带约束的优化问题，重新表示成一个无约束的函数（但引入了额外变量——乘子）

处后要求：

$\nabla_x L(x,\lambda) = 0$, $g(x) = 0$

KKT 本质上是 Lagrange Multipliers 对不等式约束的扩展

问题：为什么 Lagrange Multipliers 可以"自动处理约束"？

切空间 tangent space

对偶性 duality.

梯度 = 法向量

The basic idea is to convert a constrained problem into a form such that the derivative test of an unconstrained problem can still be applied.

Lagrangian 作为构造对偶性的核心工具

(Duality)

Duality：通过另一个视角（对偶问题）来估计原问题的最优值，并找到这个估计的上限或下限.

Lagrangian是工具，用于构造对偶问题.

Second normal form: quadratic function

~~Taylor~~ 和泰勒展开的概念不是同一回事.

level set定义：$L_c(f) = \{(x_1, \cdots, x_n) \mid f(x_1, \cdots, x_n) = c\}$

Differential topology/geometry

foundation: Smooth manifolds

Morse Lemma：微积分 → 变分法 → 拓扑.

起点：（极值与二阶近似）：为什么 critical point 附近函数"像二次型"？ → Quadratic form.

Quadratic approximations: The Taylor polynomial of order 2 generated by a twice-differentiable function $f(x)$ at $x=a$ is called the quadratic approximation of f at $x=a$.

linearization: Taylor polynomial of order 1.

题目：$f(x)=\ln(\cos x)$　$f(x)=e^{\sin x}$　$f(x)=1/\sqrt{1-x^2}$　$f(x)=\cosh x$　$f(x)=\sin x$　$f(x)=\tan x$

二阶近似本质上是用局部的二次函数（quadratic form）去替代复杂函数.

用途：判断极值类型（极小/极大/鞍点）

$f(x)=\ln(\cos x)$

$f'(x)=\frac{(\cos x)'}{\cos x}=\frac{-\sin x}{\cos x}=-\tan x$　$f'(0)=0$

$f''(x)=-(\sec^2 x)$　$f''(0)=-1$

$f(x)\approx -\frac{1}{2}x^2$