运筹学通论I

胡晓东

应用数学研究所

中国科学院数学与系统科学研究院

Http://www.amt.ac.cn/member/huxiaodong/

Institute of Applied Mathematics

Chinese Academy of Sciences

提纲

20世纪数学的五大指导理论 **Five Golden Rules**

叶其孝、刘宝光 **Great Theories of 20th Century Math**
上海教育出版社，2000 **-and Why They Matter**

1. 线性规划 对偶定理
2. 博弈论 极大极小定理
3. 非线性规划 K-K-T 定理
4. 计算/算法理论 停机定理, 库克定理

   拓扑学 不动点定理

   奇点理论 莫尔斯定理
5. 组合最优化 算法设计技巧

运筹学
- 模型
- 理论
- 算法

参考书目

Nonlinear Programming - Theory and Algorithms
Mokhtar S. Bazaraa, C. M. Shetty
John Wiley & Sons, Inc. 1979 (2nd Edition, 1993)

Linear and Nonlinear Programming
David G. Luenberger
Addison-Wesley Publishing Company, 2nd Edition, 1984/2003..

Convex Analysis **
R. T. Rockafellar
Princeton Landmarks in Mathematics and Physics, 1996.

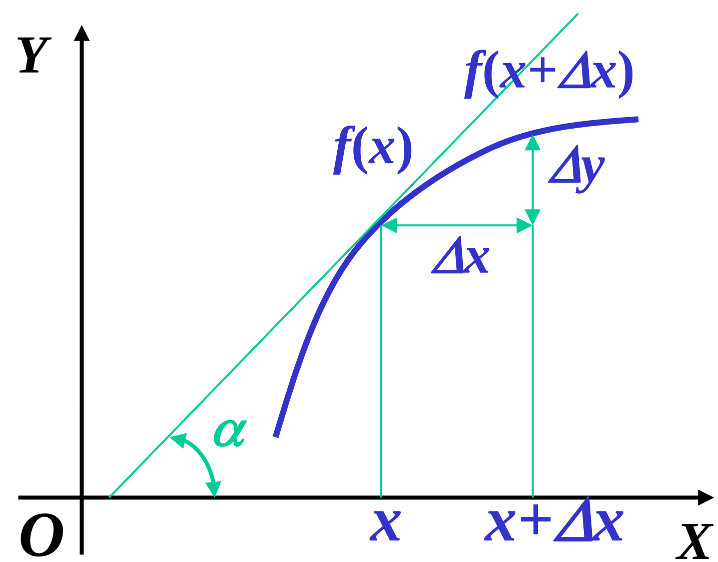
Optimization and Nonsmooth Analysis **
Frank H. Clarke
SIAM, 1990.

3. 微分复习 - 单变量函数

考虑(单变量)函数 $y = f(x)$。当自变量 x 在点 x 有一改变量Δx 时，函数 y 相应地有一改变量$\Delta y = f(x+\Delta x) - f(x)$，那么当$\Delta x$ 趋于零时，若比值$\Delta y / \Delta x$ 的极限存在，则称这个极限为函数 $f(x)$ 在点 x 的导数，记

$$y' = f'(x) = \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

从几何直观的角度看，函数 $f(x)$ 的导数 $f'(x)$ 是函数 $y = f(x)$ 表示的曲线在点 x 的切线的斜率，即 $f'(x) = \mathrm{tg}\alpha$，这里$\alpha$是曲线在点 x 处的切线与X-轴的夹角。



3. 微分复习 - 多变量函数

考虑多变量函数 $z = f(x, y)$。当变量x 在点x 有一改变量Δx，而变量 y 保持不变时，函数z 相应地有一改变量$\Delta z = f(x+\Delta x, y) - f(x, y)$，那么当$\Delta x$ 趋于零时，若比值 $\Delta z / \Delta x$ 的极限存在，则称这个极限为函数 $f(x, y)$ 在点(x, y) 关于变量 x 的偏导数，记

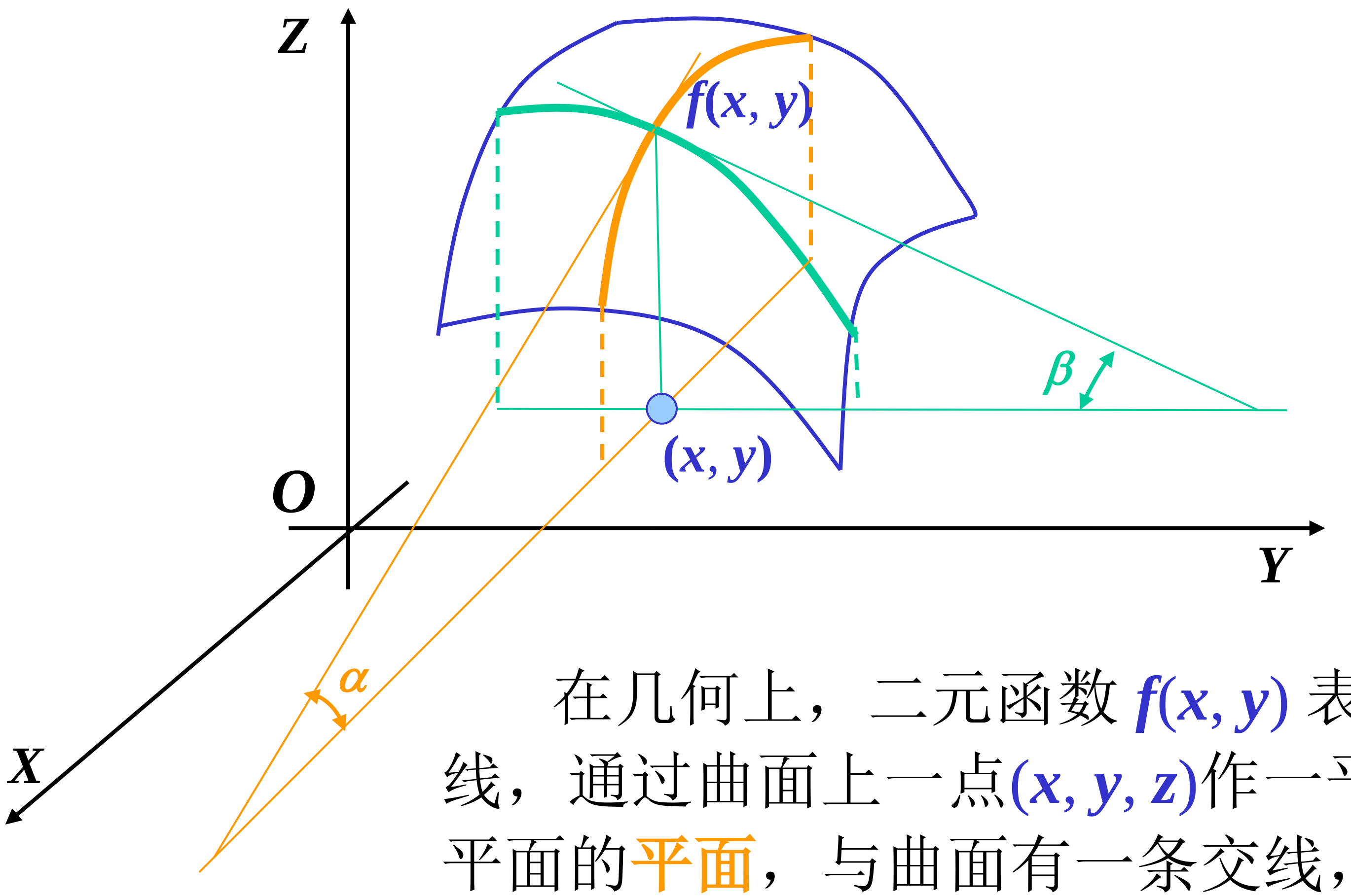
$$f'_x(x, y) = \frac{\partial f(x, y)}{\partial x} = \frac{\partial z}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x, y)}{\Delta x}$$

类似地，记多变量函数 $f(x, y)$ 在点(x, y)关于变量 y 的偏导数为

$$f'_y(x, y) = \frac{\partial f(x, y)}{\partial y} = \frac{\partial z}{\partial y} = \lim_{\Delta y \to 0} \frac{\Delta z}{\Delta y} = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x, y)}{\Delta y}$$

多变量函数的偏导数可以按照单变量函数的微分法则求出，只需要对所讨论变量求导数，其余的变量都看作常数即可。

3. 微分复习 - 多变量函数（续一）



在几何上，二元函数 ***f*(*x*, *y*)** 表示一个曲线，通过曲面上一点**(*x*, *y*, *z*)**作一平行于**OXZ**平面的**平面**，与曲面有一条交线，$\partial z/\partial x$ 就是这条曲线在该点的切线与***X***-轴正向夹角α的正切，即$\partial z/\partial x = \mathrm{tg}\alpha$。

3. 微分复习 - 多变量函数（续二）

多元函数 $y = f(x)$ 在 $x^* \in \mathrm{int}(S)$ 处是可微的，如果存在梯度 $\nabla f(x^*)$ 和一个函数$\alpha(x)$ 满足：对任意$x \in S$

$$f(x) = f(x^*) + \nabla f(x^*)^T(x-x^*) + \|x - x^*\|\alpha(x^*; x^* - x),$$

这里 $\lim_{x\to x^*} \alpha(x^*; x^* - x)=0$。事实上 $\nabla f(x)$ 的每一个分量就是相应的偏导数，即 $\nabla f(x)=(\partial f/\partial x_1, \partial f/\partial x_2, \dots, \partial f/\partial x_n)$

多元函数 $y = f(x)$ 在 $x^* \in \mathrm{int}(S)$ 处是二阶可微的，如果存在梯度 $\nabla f(x^*)$及一个海森矩阵$H(x^*)$，和一个函数$\alpha(x)$满足：$\forall x \in S$

$$f(x)=f(x^*)+\nabla f(x^*)(x-x^*)+\frac{1}{2}(x-x^*)H(x^*)(x-x^*)+\|x-x^*\|^2\alpha(x^*; x^*-x),$$

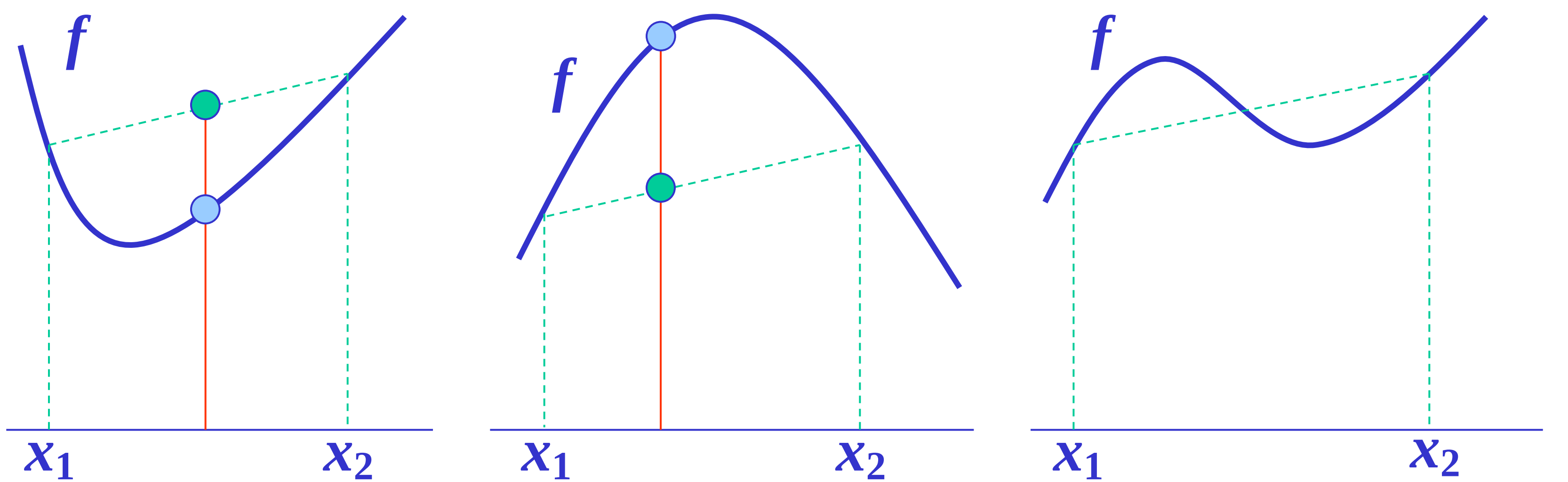
这里 $\lim_{x\to x^*} \alpha(x^*; x^*-x)=0$。事实上矩阵$H(x)$在第 i 行第 j 列的元素为$\partial^2 f/\partial x_i \partial x_j$

3. 凸分析

定义 1.称定义于凸集 S 上的函数 $f(x)$ 是**凸函数**如果对于任意 $x_1, x_2 \in S$ 和 $\lambda \in (0, 1)$，都有

$$f(\lambda x_1 + (1-\lambda)\, x_2) \le \lambda f(x_1) + (1-\lambda)\, f(x_2)。$$

称其为**严格凸**的如果不等式取严格小于号。另外，称函数 $f(x)$ 是一个 (**严格**) **凹函数** 如果 $-f(x)$ 是一个 (严格) 凸函数。

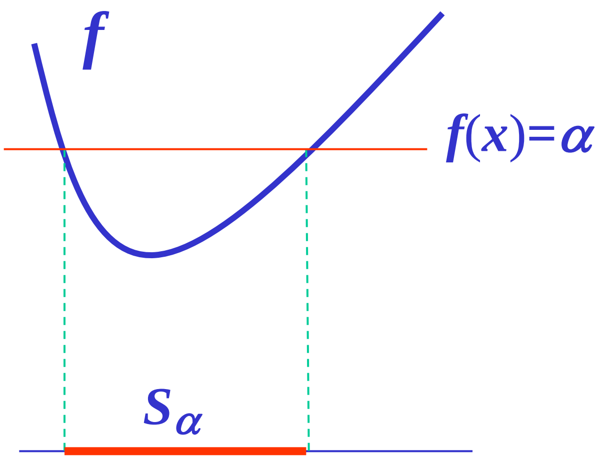


凸函数　　　　凹函数　　　　非凸非凹函数

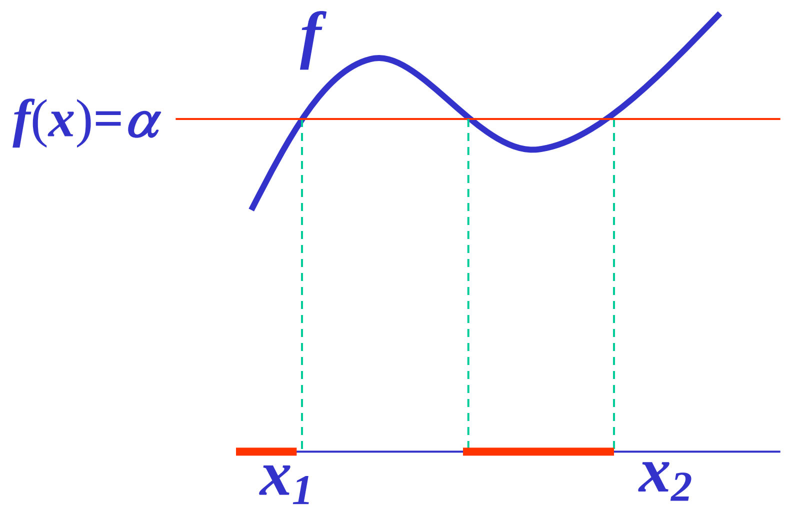
3. 凸分析 - 凸函数性质

练习. 根据凸集和凸函数的定义，证明下述两个引理。

引理 1. 设 $f(x)$ 是定义于凸集 S 上的一个凸函数。则其水平集 $S_\alpha=\{\, x\in S \mid f(x)\le \alpha \,\}$ 是一个凸集，其中 α 是任意一个实数。



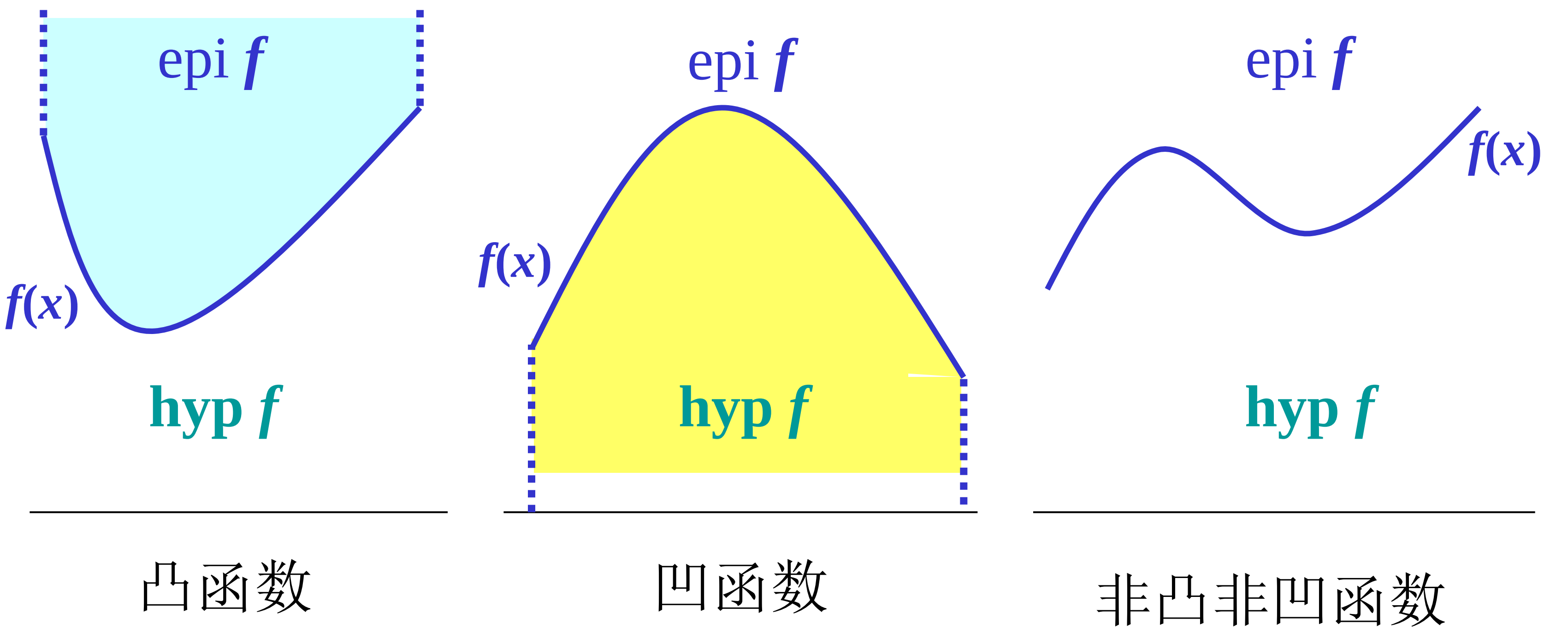
凸函数



非凸函数

3. 凸分析 - 凸函数性质（续一）

引理 2. 设 S 是n-维欧氏空间 E^n 上的一个凸集。则 $f(x)$ 是一个凸函数当且仅当其上图 $\text{epi}\, f=\{(x, y) \mid x\in S, y\in E^1, y\geq f(x)\}$ 是一个凸集。



3. 凸分析 - 凸函数性质（续二）

定理 1. 设 $f(x)$ 是定义于凸集 S 上的一个凸函数。则$f(x)$ 在集合 S 的内点集上是连续的。

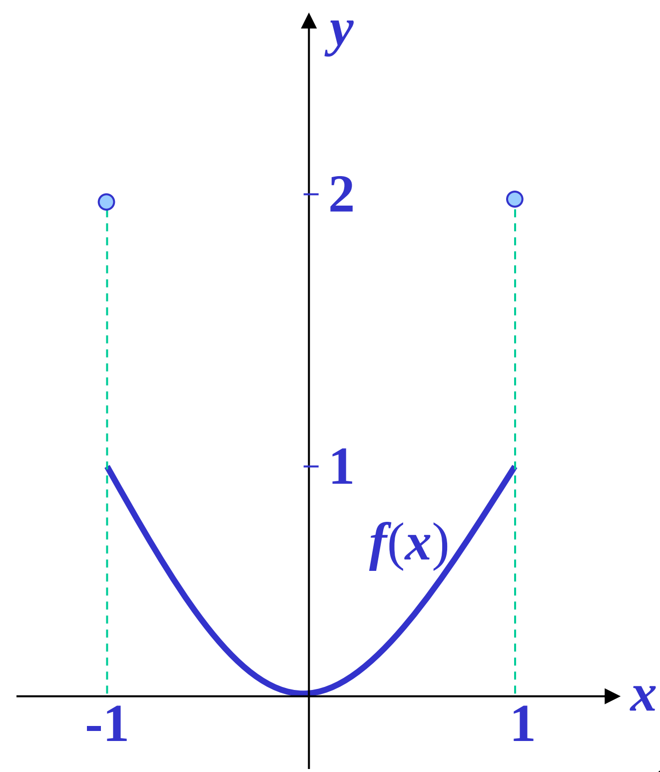
证明(练习***). 首先证明一维的情形，然后再化成一维情形。

注意，凸函数和凹函数不一定在定义域上处处连续。然而，根据上述定理，不连续性只可能发生在集合 S 的边界上。

例如，$f(x)=x^2$ 当 $|x| < 1$，

且 $f(x)= 2$ 当 $|x| = 1$，

其中 $S=\{ x \mid |x| \le 1\}$。

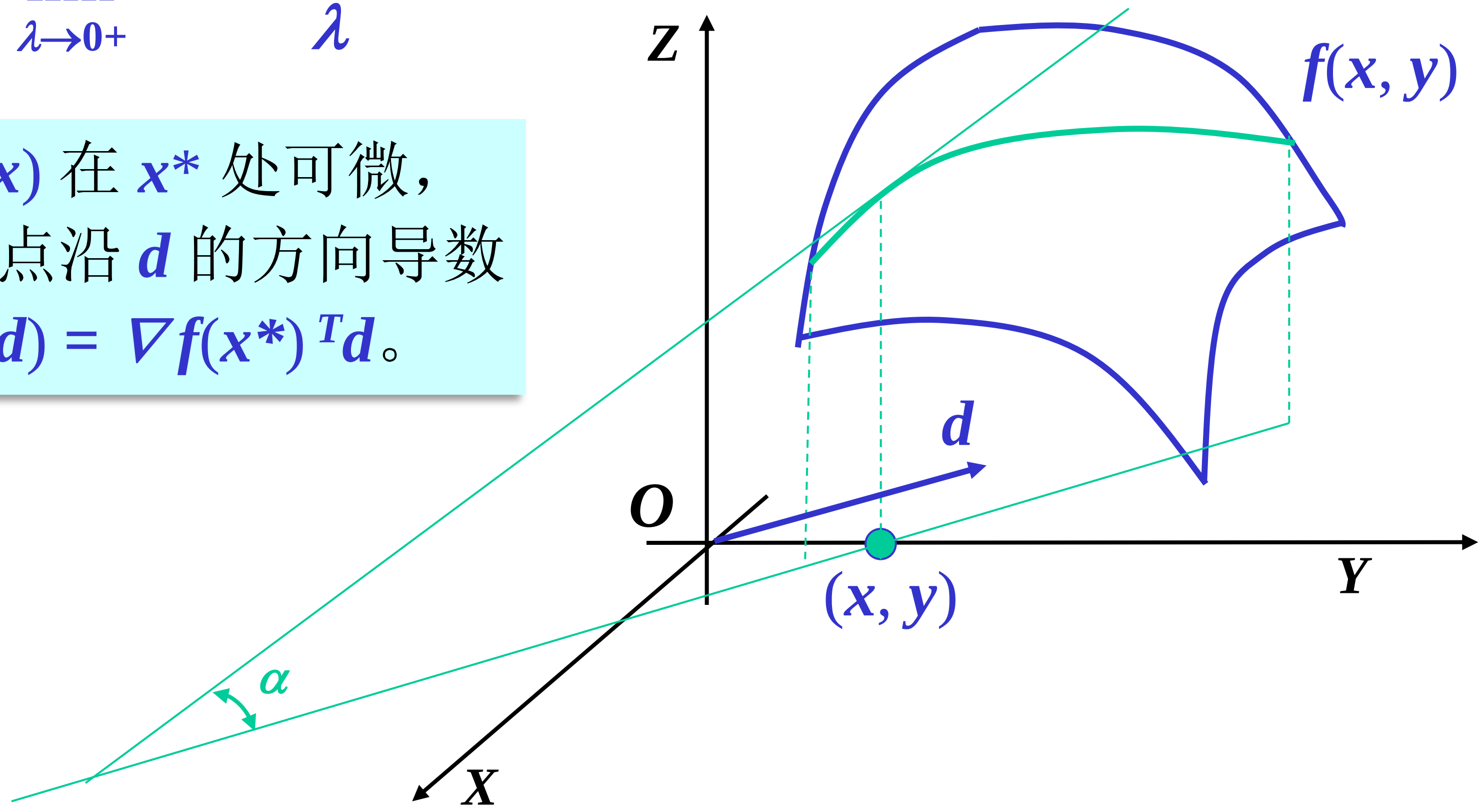


3. 凸分析 - 凸函数方向导数

定义 2. 设 $f(x)$ 是定义于集合 S 上的一个函数，另设 $x^*\in S$，d 是一个非零向量，且对足够小的 $\lambda>0$ 有 $x^*+\lambda d\in S$。函数 $f(x)$ 在 x^* 处沿 d 的方向导数为（如果极限存在）

$$f'(x^*, d) = \lim_{\lambda\to 0+} \frac{f(x^*+\lambda d) - f(x^*)}{\lambda}$$

若函数 $f(x)$ 在 x^* 处可微，则其在该点沿 d 的方向导数为 $f'(x^*, d) = \nabla f(x^*)^T d$。



3. 凸分析 - 凸函数方向导数（续一）

引理 3. 设 $f(x)$ 定义在凸集 S 上的一个凸函数，另设 $x^*\in S$，d 是一个非零向量，且满足对足够小的$\lambda > 0$，有$x^*+\lambda d\in S$。则函数 $f(x)$ 的方向导数存在。

证明 (练习*). 设 $\lambda_2 > \lambda_1 > 0$。根据函数 $f(x)$ 的凸性，我们有

$$f(x^*+\lambda_1 d) = f\left[\frac{\lambda_1}{\lambda_2}(x^*+\lambda_2 d) + \left(1-\frac{\lambda_1}{\lambda_2}\right)x^*\right]$$

$$\le \frac{\lambda_1}{\lambda_2} f(x^*+\lambda_2 d) + \left(1-\frac{\lambda_1}{\lambda_2}\right) f(x^*),$$

由此可得
$$\frac{f(x^*+\lambda_1 d) - f(x^*)}{\lambda_1} \le \frac{f(x^*+\lambda_2 d) - f(x^*)}{\lambda_2}$$

因而差商 $(f(x^*+\lambda d) - f(x^*))/\lambda$ 关于 $\lambda > 0$ 是一个单调非减的函数, 故它有极限（其值可为 ∞）。

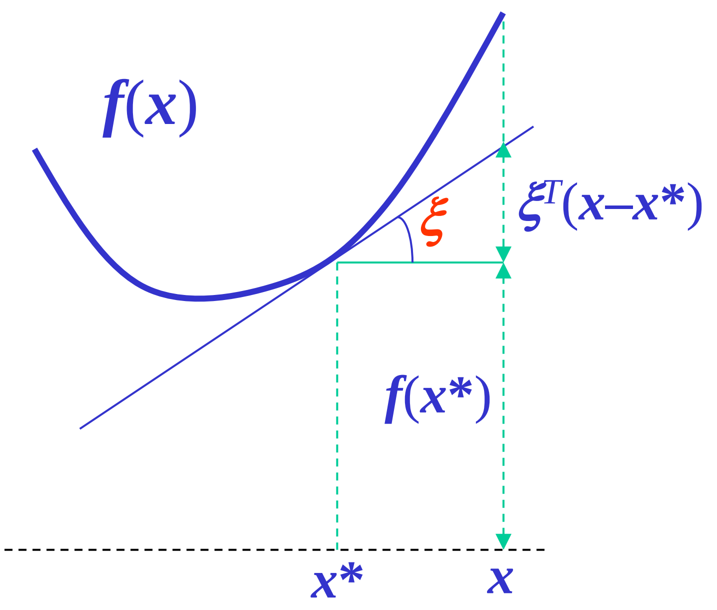
3. 凸分析 - 凸函数次梯度

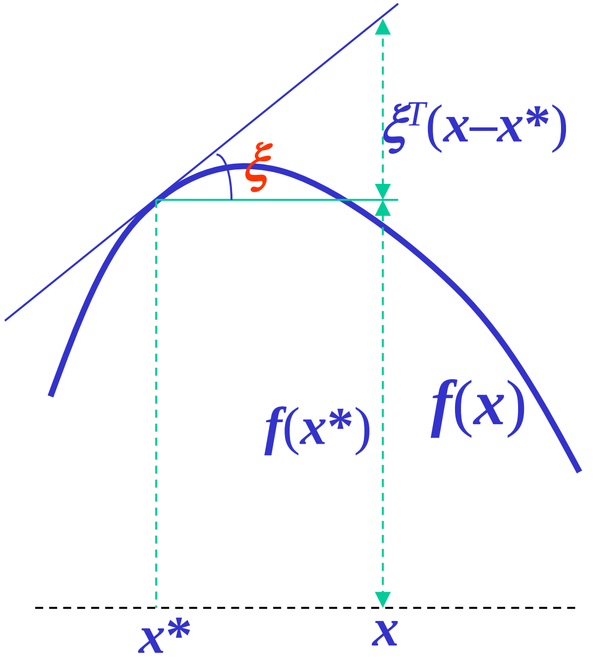
定义 3. 设 $f(x)$ 是定义于凸集 S 上的一个凸函数，另设 $x^*\in S$。称向量 ξ 为 $f(x)$ 在点 x^* 处的**次梯度**如果

$$f(x) \ge f(x^*) + \xi^T(x-x^*)，\text{对所有 } x\in S。$$

类似地，对于凹函数 $f(x)$，称 ξ 为 $f(x)$ 在点 x^* 处的**次梯度**如果

$$f(x) \le f(x^*) + \xi^T(x-x^*)，\text{对所有 } x\in S。$$





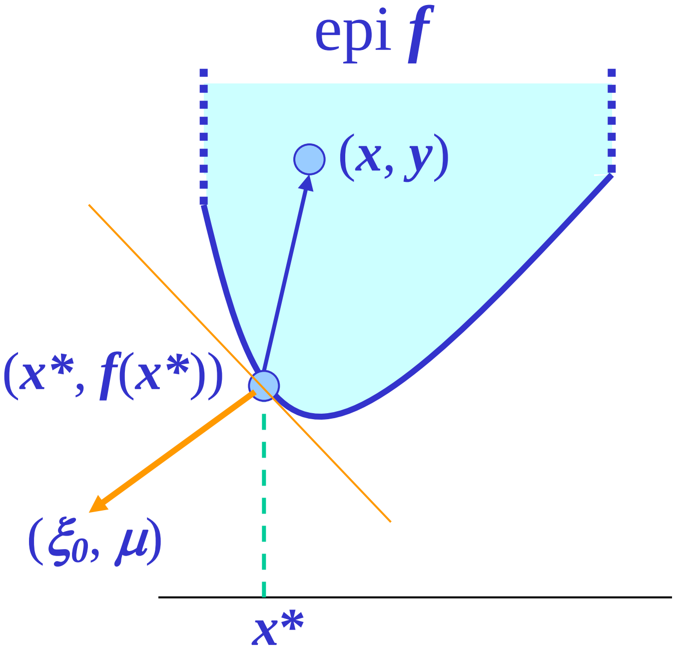
3. 凸分析 - 凸函数次梯度(续一)

定理 2. 设 $f(x)$ 是定义于凸集 S 上的一个凸函数，则 $f(x)$ 在任意一个内点 $x^*\in \mathrm{int}(S)$ 都存在次梯度。

证明 (练习).** 根据**引理 2**, **epi** f 是一个凸集。注意，$(x^*, f(x^*))$ 位于 **epi** f 的边界上，因而根据**支撑定理**，存在一个非零向量 (ξ_0, μ) 使得对所有 $(x, y)\in \mathbf{epi}\, f$ 都有

$$(\xi_0, \mu)^T(x - x^*, y - f(x^*)) \le 0$$

$$\xi_0^T(x - x^*) + \mu\,(y - f(x^*)) \le 0$$



再注意，我们有 $\mu \le 0$（否则我们可以取任意大的 y, ……）；事实上，$\mu < 0$（否则有 $\xi_0^T(x - x^*) \le 0$，然而我们可以取充分小的 $\lambda > 0$ 使得 $x^*+ \lambda\xi_0 \in \mathrm{int}(S)$, ……）

3. 凸分析 - 凸函数次梯度(续二)

定理 3. 设 S 是 E^n 上的一个凸集，函数 $f(x)$：$S \to E^1$。若对每一个内点 $x^* \in \text{int}(S)$ 都存在一个次梯度向量 ξ 使得，对所有 $x \in S$，都有 $f(x) \geq f(x^*) + \xi(x - x^*)$。则函数 $f(x)$ 在内点集 $\text{int}(S)$ 上是凸的。

证明 (练习*). 分别取 $x = x_1$，x_2，$x^* = \lambda x_1 + (1-\lambda) x_2$，则有

$$f(x_1) \geq f(\lambda x_1 + (1-\lambda) x_2) + (1-\lambda)\,\xi(x_1 - x_2)，\text{和}$$
$$f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2) + \lambda\,\xi(x_2 - x_1)。$$

将上述两个不等式分别乘以 λ 和 $(1-\lambda)$，再相加，即可得

$$\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2).$$

3. 凸分析 - 凸函数梯度

定理 4. 设 S 是 E^n 上的一个非空凸的开集，函数 $f(x)$: $S \to E^1$。若 $f(x)$ 在点 x^* 处可微，则 $f(x)$ 在点 x^* 处的次梯度向量集合恰好含有一个向量 $\nabla f(x^*)$。

证明. 设 ξ 是函数 $f(x)$ 在点 x^* 处的一个次梯度向量。则有

$$f(x^*+\lambda d) \ge f(x^*) + \lambda \xi^T d,$$
$$f(x^*+\lambda d) = f(x^*) + \lambda \nabla f(x^*) d + \lambda \|d\| \alpha(x^*; \lambda d),$$

将上述等式带入不等式，即得

$$0 \ge \lambda (\xi - \nabla f(x^*))^T d - \lambda \|d\| \alpha(x^*; \lambda d).$$

如果将上述不等式两边同除以 $\lambda > 0$，并让 $\lambda \to 0$，那么即可得到 $0 \ge (\xi - \nabla f(x^*))^T d$。现取 $d = \xi - \nabla f(x^*)$，即得 $\xi = \nabla f(x^*)$。

3. 凸分析 - 凸函数梯度(续一)

定理 5. 设 S 是 E^n 上的一个非空凸的开集，函数 $f(x)$： $S \to E^1$ 是可微的。则 $f(x)$ 是一个凸函数当且仅当对任意 $x^* \in S$，都有

$$f(x) \geq f(x^*) + \nabla f(x^*)(x - x^*)，所有\ x \in S。$$

类似地， $f(x)$ 是一个严格凸的当且仅当对每一个 $x^* \in S$，都有

$$f(x) > f(x^*) + \nabla f(x^*)(x - x^*)，所有\ x \in S。$$

定理 6. 设 S 是 E^n 上的一个非空凸的开集，函数 $f(x)$： $S \to E^1$ 是可微的。则 $f(x)$ 是一个凸函数当且仅当对任意 $x_1, x_2 \in S$，都有

$$(\nabla f(x_2) - \nabla f(x_1))^T (x_2 - x_1) \geq 0。$$

类似地，函数 $f(x)$ 是严格凸的当且仅当对任意不同的两个向量 $x_1, x_2 \in S$，都有

$$(\nabla f(x_2) - \nabla f(x_1))^T (x_2 - x_1) > 0。$$

3. 凸分析 - 凸函数梯度(续二)

证明. "当"： 根据**定理 5**，对任意 $x_1, x_2 \in S$， 都有。

$$f(x_1) \geq f(x_2) + \nabla f(x_2)(x_1 - x_2) \text{ 和}$$

$$f(x_2) \geq f(x_1) + \nabla f(x_1)(x_2 - x_1)\text{。}$$

将上述两个不等式相加， 即得 $(\nabla f(x_2) - \nabla f(x_1))(x_2 - x_1) \geq 0$。

"仅当"： 根据**中值定理**， 可得

$f(x_2) - f(x_1) = \nabla f(x)(x_2 - x_1)$, 其中$x = \lambda x_1 + (1 - \lambda)x_2$， $\lambda \in (0, 1)$。

再由假设 $(\nabla f(x) - \nabla f(x_1))^T(x - x_1) \geq 0$, 可得

$$(1 - \lambda)(\nabla f(x) - \nabla f(x_1))^T(x_2 - x_1) \geq 0\text{。}$$

因而有 $\nabla f(x)(x_2 - x_1) \geq \nabla f(x_1)(x_2 - x_1)$， 由此可得

$$f(x_2) \geq f(x_1) + \nabla f(x_1)(x_2 - x_1).$$

再根据**定理 5**， 即知 $f(x)$ 是一个凸函数。

3. 凸分析 - 凸函数二阶微分

定理 7. 设 S 是 E^n 上的一个非空凸的开集， $f(x)$： $S \to E^1$ 是一个二次可微函数。 则 $f(x)$ 是一个凸函数当且仅当 S 中的每一个点处的海森矩阵都是半正定的。

证明 (练习).** $\Rightarrow$ 根据**定理 5** 和二次可微函数的定义，我们有

$$f(x^*+\lambda x) \geq f(x^*) + \lambda \nabla f(x^*)(x) \text{ 和}$$

$$f(x^*+\lambda x) = f(x^*) + \lambda \nabla f(x^*) x + 0.5\, \lambda^2 x^T H(x^*) x + \lambda^2 \|x\|^2 \alpha(x^*; \lambda x)$$

将上述等式带入不等式，即得 $0.5\lambda^2 x^T H(x^*) x + \lambda^2 \|x\|^2 \alpha(x^*; \lambda x)$。现令 $\lambda \to 0$，既有 $x^T H(x^*) x \geq 0$。

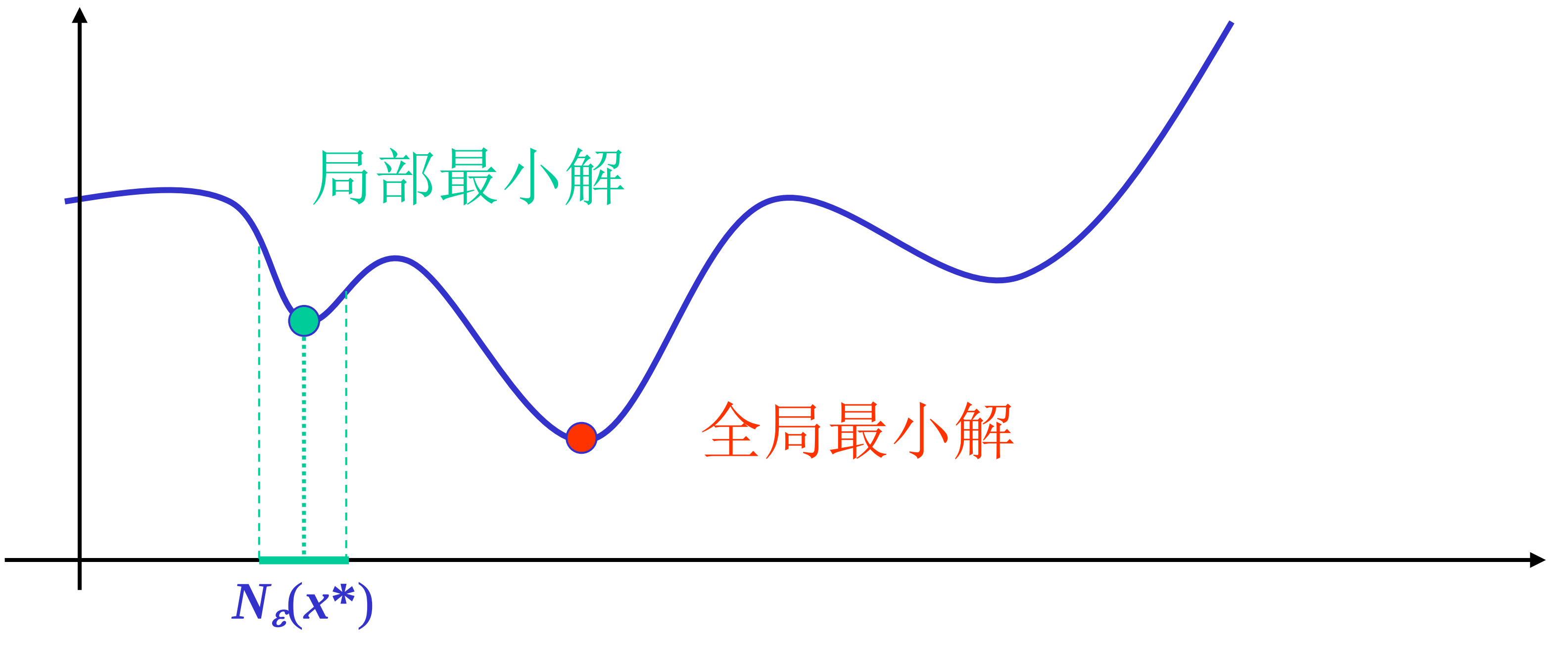
$\Leftarrow$ 根据**中值定理**，即可知，存在 $x' = \lambda x^* + (1-\lambda)x$ 使得

$$f(x) = f(x^*) + \nabla f(x^*)(x - x^*) + 0.5(x - x^*)^T H(x')(x - x^*)$$

又因 $H(x')$ 是半正定矩阵，故有 $f(x) \geq f(x^*) + \nabla f(x^*)(x - x^*)$, ……

3. 凸分析 - 凸函数最优条件

定义 5. 考虑求最小值问题 $\min\{f(x) \mid x \in S \subset E^n\}$。设 $x^* \in E^n$。若对所有 $x \in S$ 都有 $f(x^*) \le f(x)$，则称 $x^* \in S$ 是一个**全局最小解**或者**全局最优解**。 若存在 x^* 的一个邻域 $N_\varepsilon(x^*)$，使得对任意 $x \in S \cap N_\varepsilon(x^*)$，都有 $f(x^*) \le f(x)$，则称 x^* 为一个**局部最小解**或者**局部最优解**。



3. 凸分析 - 凸函数最优条件

定理 8. 设 S 是 E^n 上的一个非空凸集，函数 $f(x)$：$S \to E^1$。另设 $x^* \in S$ 是求最小值问题 $\min\{ f(x) \mid x \in S \}$ 的一个局部最优解。

i) 若 $f(x)$ 是一个凸函数，则 x^* 是一个全局最优解。

ii) 若 $f(x)$ 是一个严格凸函数，则 x^*是惟一一个全局最优解。

证明. **(练习)** (i) 采用反证法。假设 $f(x^*) > f(x')$，其中 $x' \in S$。则根据函数 $f(x)$ 的凸性可知，对任意一个 $\lambda \in (0, 1)$，都有

$$f(\lambda x' + (1-\lambda)\, x^*) \le \lambda f(x') + (1-\lambda) f(x^*)$$
$$< \lambda f(x^*) + (1-\lambda) f(x^*) = f(x^*), \ldots$$

(ii) 还是采用反证法。假设存在另外一个全局最优解 $x' \neq x^*$，$f(x') = f(x^*)$。根据函数 $f(x)$ 的严格凸性，我们有

$$f(x'/2 + x^*/2) < f(x')/2 + f(x^*)/2 = f(x^*), \ldots\ldots$$

3. 凸分析 - 凸函数最优条件（续一）

定理 9. 设函数 $f(x)$： $S\to E^1$是一个凸函数，其中S 是一个非空凸集。则 $x^*\in S$ 是求最小值问题 $\min\{ f(x) \mid x\in S \}$的全局最优解当且仅当 $f(x)$ 在 x^* 处的一个次梯度 ξ 满足，对任意 $x\in S$，都有 $\xi^T(x-x^*)\geq 0$。

证明. (练习)** $\Leftarrow$ 根据函数 $f(x)$ 的凸性和次梯度的性质，…

$\Rightarrow$ 根据在 x^* 处次梯度ξ 的定义。定义 E^{n+1} 中的两个集合：

$$S_1=\{(x-x^*, y) \mid x\in E^n, y>f(x)-f(x^*)\},$$

$$S_2=\{(x-x^*, y) \mid x\in S, y\leq 0\}。$$

因为 x^* 是一个最优解，则有 $S_1\cap S_2=\varnothing$。再根据**分离定理**，可得 $\inf\{px \mid x\in S_1\}\geq \sup\{ px \mid x\in S_2\}$。因此存在一个非零向量 (ξ_0, μ) 和一个实数 α 满足

$$\xi_0 (x-x^*)+\mu y\leq \alpha,\ \ x\in E^n,\ \ y>f(x)-f(x^*),$$

$$\xi_0 (x-x^*)+\mu y\geq \alpha,\ \ x\in S,\quad y\leq 0。$$

3. 凸分析 - 凸函数最优条件（续二）

证明.（续前）

$$\xi_0^T(x-x^*)+\mu y\le\alpha,\ x\in E^n,\ y>f(x)-f(x^*),\qquad\text{(i)}$$

$$\xi_0^T(x-x^*)+\mu y\ge\alpha,\ x\in S,\quad y\le 0。\qquad\text{(ii)}$$

在 (ii) 中令 $x=x^*$， $y=0$， 即得 $\alpha\le 0$。再在(i) 中令 $x=x^*$，既有 $\mu\le 0$ 和 $\alpha\ge 0$ (因为 y 可取任意大或小)。因此有 $\alpha=0$。

当 $\mu=0$ 时， 在 (i)中令 $x=x^*+\xi_0$， 则有 $\xi_0^T(x-x^*)=\|\xi_0\|^2\le 0$。然而， $(\xi_0,\mu)\ne 0$， 故有 $\mu<0$。现用$-\mu$ 同除 (i) 和 (ii)， 即得

$$\xi^T(x-x^*)\le y,\ x\in E^n,\quad y>f(x)-f(x^*),\qquad\text{(iii)}$$

$$\xi^T(x-x^*)-y\ge 0,\quad x\in S,\quad y\le 0\ 。\qquad\text{(iv)}$$

在 (iv)式中， 令 $y=0$， 即得$\xi^T(x-x^*)\ge 0$。再根据 (iii)式， 且令 $y\to f(x)-f(x^*)$， 即得 $f(x)\ge f(x^*)+\xi^T(x-x^*)$。

3. 凸分析 - 凸函数最优条件（续三）

定理 10. 设函数 $f(x)$: $S \to E^1$是一个凸函数，其中S是一个非空凸集。若 S 是开集，则 $x^* \in S$ 是求最小值问题 $\min\{f(x) \mid x \in S\}$ 的一个最优解当且仅当 $f(x)$ 在 x^* 处 $\mathbf{0}$ 向量是其次梯度。

证明. (练习) 由**定理 9** 可知，对任意 $x \in S$ 都有$\xi^T(x - x^*) \geq 0$。令 $x = x^* - \lambda\xi \in S$，其中 $\lambda > 0$。则有 … …

定理 11. 设函数 $f(x)$: $S \to E^1$ 是一个凸函数，其中S是一个非空凸集。 若 $f(x)$ 是可微的，则 $x^* \in S$ 是求最小值问题 $\min\{ f(x) \mid x \in S \}$ 的一个最优解当且仅当，对所有 $x \in S$，都有 $\nabla f(x^*)(x - x^*) \geq 0$。而且，若 S 是开集，则 $x^* \in S$ 是一个最优解当且仅当 $\nabla f(x^*) = \mathbf{0}$。

3. 凸分析 - 凸函数最优条件（续四）

下面我们考虑求定义在凸集上的凸函数的最大值问题。与求最小值问题不同，我们不能根据一个解的某些局部信息，找到一个更好的解。因此，求凸函数的最大值问题通常要比求凸函数的最小值问题困难得多。

定理 12. 设函数 $f(x)$：$S\to E^1$ 是一个凸函数，其中 S 是一个非空凸集。若 $x^*\in S$ 是求最大值问题 $\max\{ f(x) \mid x\in S\}$ 的一个局部最优解，则对所有 $x\in S$, 都有 $\xi^T(x - x^*) \le 0$，其中 ξ 是 $f(x)$ 在点 x^* 处的任一次梯度。

证明. (练习) 当 $\lambda > 0$ 充分小时，有 $x^*+\lambda(x - x^*)\in S$，且有 $f(x^*+\lambda(x - x^*)) \le f(x^*)$。对函数 $f(x)$ 在点 x^* 处的任一次梯度 ξ, 都有 $f(x^*+ \lambda(x - x^*)) - f(x^*) \ge \lambda \xi^T(x - x^*)$。因而 ……

3. 凸分析 - 凸函数最优条件（续五）

定理 13. 设函数 $f(x)$：$S \to E^1$ 是一个可微的凸函数，其中$S \subset E^n$是一个非空凸集。若 $x^* \in S$ 是求最大值问题 $\max\{ f(x) \mid x \in S\}$ 的一个局部最优解，则对所有$x \in S$，都有 $\nabla f(x^*)(x - x^*) \le 0$。

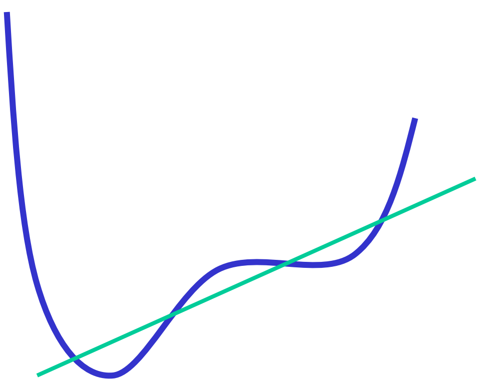
定理 14. 设函数 $f(x)$：$S \to E^1$ 是一个凸函数，其中$S \subset E^n$是一个非空的紧的多面体。则求最大值问题 $\max\{f(x) \mid x \in S\}$存在一个最优解 x^*，它是 S 的一个极点。

证明 (练习). 在定理的假设条件下，$f(x)$有一个最优解 $x^* \in S$。根据多面体的表示定理，我们有

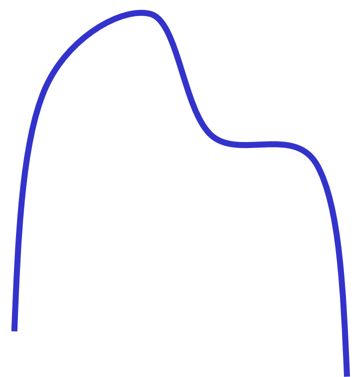
$$x^* = \sum \lambda_i x_i, \ \ldots\ldots$$

3. 凸分析 - 凸函数的推广

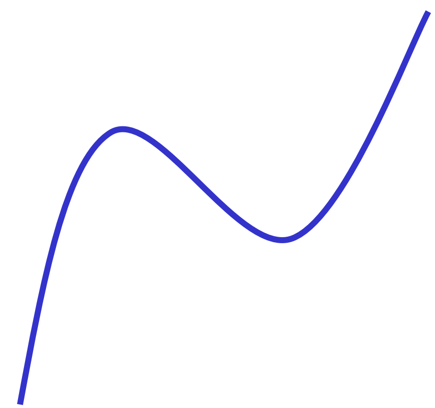
Definition 6. Let $f(x)$ be a function defined on a nonempty convex set S. It is called a **quasiconvex function** if for each pair of $x_1 \in S$, $x_2 \in S$, $f(\lambda x_1+(1-\lambda)x_2) \le \max\{f(x_1), f(x_2)\}$ for each $\lambda \in (0, 1)$. The function $f(x)$ is said to be **quasiconcave** if $-f(x)$ is quasiconvex.



quasiconvex



quasiconcave



neither quasiconvex nor quasiconcave

3. 凸分析 - 凸函数的推广（续一）

Lemma 4. Let $f(x)$ be a function defined on a nonempty convex set S in E^n. Then $f(x)$ is quasiconvex if and only if for each real number α, $S_\alpha=\{ x\in S \mid f(x) \le \alpha \}$ is a convex set.

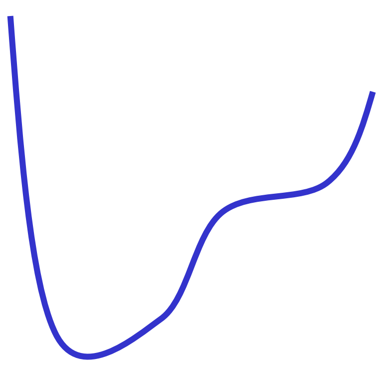
Proof. (**Exercise**). By definitions of convexity and quasiconvexity.

Theorem 15. Let $f(x)$ be a quasiconvex function defined on a nonempty compact polyhedral set S in E^n. If f is continuous on S. Then there exists an optimal solution x^* to the problem of maximizing $f(x)$ subject to $x\in S$ that is an extreme point of S.

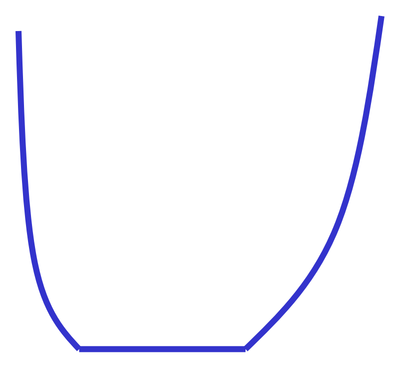
Proof (**Exercise**). Under the assumptions, $f(x)$ has a maximum at $x^*\in S$. By **Representation Theorem**, we have $x^*= \sum\lambda_i x_i$. By contradiction argument, consider $S_\alpha= \{x\in S \mid f(x) \le\alpha \}$, where $\alpha = \max f(x_i)$

3. 凸分析 - 凸函数的推广（续二）

Definition 7. Let $f(x)$ be a function defined on a nonempty convex set S. It is called a **strictly quasiconvex function** if for each pair of $x_1, x_2 \in S$ with $f(x_1) \neq f(x_2)$, $f(\lambda x_1 + (1-\lambda)x_2) < \max\{f(x_1), f(x_2)\}$ for each $\lambda \in (0, 1)$. The $f(x)$ is said to be **strictly quasiconcave** if $-f(x)$ is strictly quasiconvex.



Strictly quasiconvex



Strictly quasiconvex

Quasiconvex but not strictly quasiconvex

3. 凸分析 - 凸函数的推广（续三）

Exercise. Every strictly convex function is a convex function. But some strictly quasiconvex functions are not quasiconvex. Here is a simple example: $f(x)=1$ if $x=0$, and $f(x)=0$ if $x\neq 0$.

Definition 8. Let $f(x)$ be a function defined on a nonempty set S. It is called a **lower semicontinuous function** at $x^*\in S$ if for any $\varepsilon > 0$, there exists a $\delta>0$ such that $x\in S$ and $\|x - x^*\| <0$ imply $f(x) - f(x^*) > -\varepsilon$. Similarly, $f(x)$ is called a **upper semicontinuous function** at $x^*\in S$ if for each $\varepsilon>0$, there exists a $\delta>0$ such that $x\in S$ and $\|x - x^*\| < 0$ imply $f(x) - f(x^*) < \varepsilon$.

Exercise**. (1) If $f(x)$ is lower (upper) semicontinuous, then it achieves a minimum (maximum) over a nonempty compact set.
(2) A continuous function is both lower and upper semicontinuous.

3. 凸分析 - 凸函数的推广（续四）

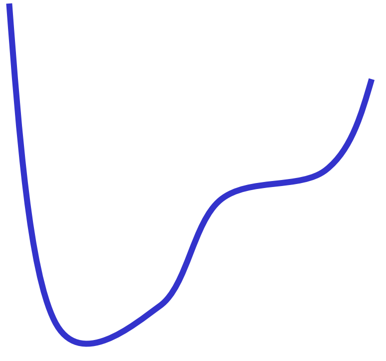
Lemma 3. Let $f(x)$ be a strictly quasiconvex function defined on a nonempty convex set S in E^n. If $f(x)$ is lower semicontinuous, then $f(x)$ is quasiconvex.

Theorem 16. Let $f(x)$ be a strictly quasiconvex function defined on a nonempty convex set S in E^n. If x^* is a local optimal solution to the problem of minimizing $f(x)$ subject to $x \in S$, then it is also a global optimal solution.

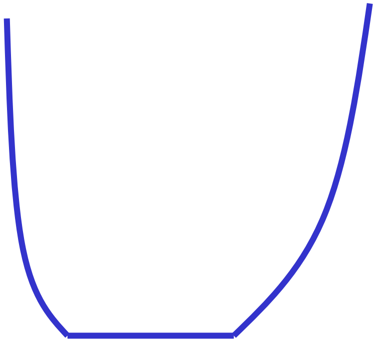
Proof (Exercise). Suppose, by contradiction, that $f(x') < f(x^*)$. Since x^* is a local optimal solution, $f(x^*) \le f(\lambda x' + (1-\lambda) x^*)$ for sufficiently small $\lambda > 0$, and $f(x)$ is a strictly quasiconvex function, $f(\lambda x' + (1-\lambda) x^*) < \max\{ f(x'), f(x^*) \} = f(x^*)$.

3. 凸分析 - 凸函数的推广（续五）

Definition 9. Let $f(x)$ be a function defined on a nonempty convex set S. It is called a **strongly quasiconvex function** if for each distinct $x_1, x_2 \in S$, $f(\lambda x_1+(1-\lambda)x_2) < \max\{ f(x_1), f(x_2) \}$ for $\lambda \in (0, 1)$. The function $f(x)$ is said to be **strongly quasiconcave** if $-f(x)$ is strongly quasiconvex.



strictly and strong quasiconvex



strictly but not strong quasiconvex

3. 凸分析 - 凸函数的推广（续六）

Exercise* Prove the following three statements:

(i) Every strictly convex function is strongly quasiconvex.

(ii) Every strongly quasiconvex function is strictly quasiconvex.

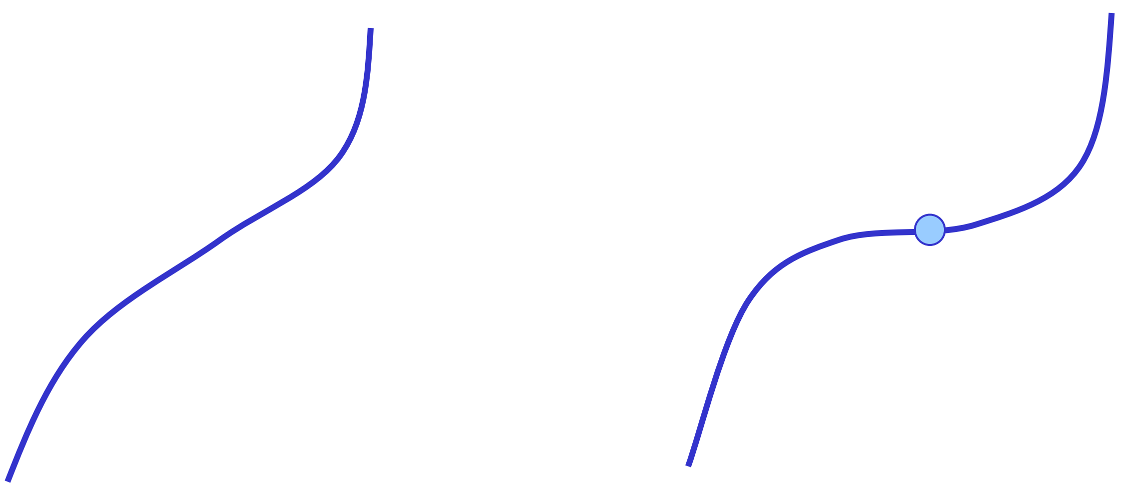
(iii) Every strongly quasiconvex function is quasiconvex.

Theorem 17. Let $f(x)$ be a strongly quasiconvex function defined on a nonempty convex set S in E^n. If x^* is a local optimal solution to the problem of minimizing $f(x)$ subject to $x \in S$, then it is the unique global optimal solution.

Proof (Exercise). Suppose, by contradiction, that $f(x') \leq f(x^*)$ and $x' \neq x^*$. Since x^* is a local optimal solution, $f(x^*) \leq f(\lambda x' + (1-\lambda)x^*)$ for sufficiently small $\lambda > 0$, and f is a strongly quasiconvex function, $f(\lambda x' + (1-\lambda)x^*) < \max\{f(x'), f(x^*)\} = f(x^*)$.

3. 凸分析 - 凸函数的推广（续七）

Definition 10. Let $f(x)$ be a differentiable function defined on a nonempty open set S. It is called **pseudoconvex** if for each pair of $x_1 \in S$, $x_2 \in S$ with $\nabla f(x_1)(x_2 - x_1) \geq 0$, we have $f(x_2) \geq f(x_1)$. The function $f(x)$ is said to be **pseudoconcave** if $-f(x)$ is pseudoconvex.



pseudoconvex

not pseudoconvex but strictly quasiconvex

3. 凸分析 - 凸函数的推广（续八）

Exercise. Suppose that function $f(x)$ is pseudoconvex. If $\nabla f(x^*)=0$, then x^* is a global minimum of $f(x)$.

Theorem 18. Suppose that $f(x)$ is a differentiable pseudoconvex function defined on a nonempty open convex set S in E^n. Then $f(x)$ is both strictly quasiconvex and quasiconvex.

Definition 11. Let $f(x)$ be a differentiable function defined on a nonempty open set S. It is called **strictly pseudoconvex** if for each distinct $x_1 \in S$, $x_2 \in S$ with $\nabla f(x_1)(x_2 - x_1) \geq 0$, we have $f(x_2) > f(x_1)$.

Theorem 19. Suppose that $f(x)$ is a differentiable strictly pseudoconvex function defined on a nonempty open convex set S in E^n. Then $f(x)$ is strongly quasiconvex.

3. 凸分析 - 凸函数的推广（续九）

Relationship among various types of convexity