lec43.

(测验)证明: $\max\limits_{v,\lambda_1,\lambda_2} -b^Tv-\lambda_1^Tu+\lambda_2^Tl.$ 与 $\max -b^Tv+l^T(A^Tv+l)^+ - u^T(A^Tv+l)^-$ 等价

$\text{s.t.}\quad A^Tv+\lambda_1-\lambda_2+c=0.$

$\lambda_1\geq 0,\ \lambda_2\geq 0.$

证明: 若 $g(x,u,w)$ 为凸, 则 $p(u,w)=\inf\limits_{x} g(x,u,w)$ 为凸.

无约束优化问题. $\min f(x)$.

迭代算法. $\underline{\underline{x}}^{k+1}=x^k+\underline{\underline{\alpha}}^k\underline{\underline{d}}^k.$

(下降算法).

$\alpha^k=\arg\min\limits_{\alpha} f(x^k+\alpha^kd^k)$

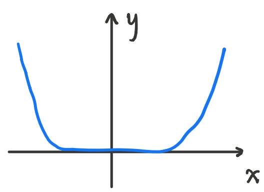
$\min f(x)$ $\underline{f(x)\text{一阶可微}}$. 最优性条件.

$\nabla f(x)=0.$ $\nabla f(\tilde{x})\approx 0$, $x\to x^*$? $f(x)\to f(x^*)$.

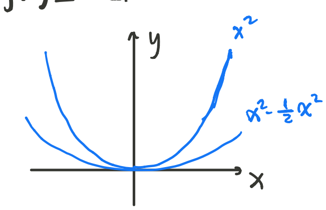
lec44.

假设 $f(x)$二阶可微且有强凸性.

$\exists m>0,\ \forall x\in \mathrm{dom} f,\ \nabla^2 f(x)\succeq mI.$



不是强凸的.



是强凸的.

$\forall x,y\in \mathrm{dom} f\quad f(y)\geq f(x)+\nabla f^T(x)(y-x)+\frac{1}{2}m\|y-x\|_2^2$

当$\nabla f(x)\to 0$时, $f(x)\to f(x^*)$?

x给定, $f(x)+\nabla f^T(x)(y-x)+\frac{1}{2}m\|y-x\|_2^2$. 是$y$的凸函数.

$\nabla f^T(x)+m(\tilde{y}-x)=0\quad \tilde{y}=x-\frac{\nabla f^T(x)}{m}.$

$\geq f(x)-\frac{1}{2m}\|\nabla f(x)\|_2^2.$

$p^*\geq f(x)-\frac{1}{2m}\|\nabla f(x)\|_2^2$

$\underline{p^*+\frac{1}{2m}\|\nabla f(x)\|_2^2\geq f(x)\geq p^*}.$

$$\|f(x)-p^*\|_2 \le \frac{1}{2m}\|\nabla f(x)\|_2^2.$$

当 $\nabla f(x)\to 0$ 时，$x\to x^*$?

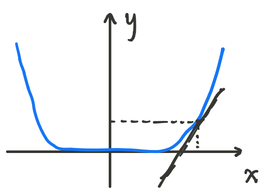
$$f(x)\ge p^* = f(x^*) \ge f(x)+\nabla f(x)(x^*-x)+\frac{m}{2}\|x^*-x\|_2^2.$$

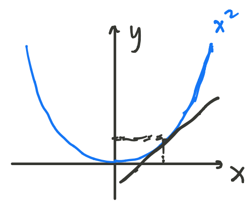
$\langle a,b\rangle+\|a\|\|b\|\ge 0.$

$$\ge f(x)-\|\nabla f(x)\|_2\|x^*-x\|_2+\frac{m}{2}\|x^*-x\|_2^2.$$

$$f(x)\ge p^* \Rightarrow -\|\nabla f(x)\|_2\|x^*-x\|_2+\frac{m}{2}\|x^*-x\|_2=0.$$

$$\|x^*-x\|_2\le\frac{2}{m}\|\nabla f(x)\|_2.$$





$$\exists M>0,\ \forall x\in \mathrm{dom} f,\ \nabla^2 f(x)\preceq MI.$$

$$\forall x,y\in \mathrm{dom} f,\ f(y)\le f(x)+\nabla f^T(x)(y-x)+\frac{M}{2}\|y-x\|_2^2$$

$$p^*\le f(x)-\frac{1}{2M}\|\nabla f(x)\|_2^2.$$

梯度大，离最优值远

梯度下降法. $d^k=-\nabla f(x^k).$

Repeat $d^k=\underset{0\le\alpha\le\alpha_{max}}{\arg\min} f(x^k+\alpha d^k).$ $\longrightarrow$ $\begin{cases}\text{exact}\\ \text{inexact.}\end{cases}$

$$x^{k+1}=x^k+\alpha^k d^k$$

Until Convergence.