Conlex Optimization 12/18/2021

1. 集合

$\theta_1x_1+\dots+\theta_nx_n\in C$

凸集: $\forall x_1,x_2\in C$, $0\le\theta\le1$, $\theta x_1+(1-\theta)x_2\in C$

仿射集合: $\forall x_1,x_2,\theta$, $\theta x_1+(1-\theta)x_2\in C$, C仿射

锥: $\forall x_1,x_2$, $\theta\ge0$, … $\in C$, C锥

集合定义: 内(相对)内部: 至一个以x为中心的球包含于C, int C; $\{\|y-x\|_2\le\epsilon\}\subseteq C$

闭包: $\forall\epsilon>0, \exists y\in C, \|x-y\|_2\le\epsilon$

边界: cl C \ int C

相对内部 relint: 仿射包 aff(C)的内部

凸包: $\forall x_1,x_2\in C$, 凸包 conv C

仿射包: 仿射组合全体

2. 凸集例子

① 超平面 $\{x|a^Tx=b\}$; 半空间 $\{x|a^Tx\le b\}$, $\{x|a^Tx\ge b\}$

② 球 $(x-x_c)^T(x-x_c)\le r^2$; 椭球 $(x-x_c)^TP^{-1}(x-x_c)\le1$ (P正定)

③ 范数锥 $\{(x,t)\,|\,\|x\|\le t\}$

④ 多面体: 有限个半空间/超平面交集

单纯形: $v_0\ldots v_k$ 仿射独立, 凸包 $C=\{\theta_0v_0+\dots+\theta_kv_k\,|\,\theta\succeq0, 1^T\theta=1\}$

⑤ 半正定锥: S^n_+ 是一个凸锥

3. 保凸运算 (集合)

① 交集: 有限/无穷皆可

② 仿射函数 $f(x)=Ax+b$, $S\to f(S)$, S凸 ⇒ f(S)凸

③ 透视函数 $P: R^{n+1}\to R^n$, $P(z,t)=z/t$; $C\subseteq\text{dom}\,P$ ⇒ P(C)凸

线性分式: $f(x)=\frac{Ax+b}{c^Tx+d}$, $\text{dom}\,f=\{x|c^Tx+d>0\}$

① 超平面分离定理: 若C, D凸, 且 $C\cap D=\emptyset$, 则 $\exists a\ne0, b$, 使 $\forall x\in C, a^Tx\le b$; $\forall x\in D, a^Tx\ge b$. $a^Tx=b$: 分离超平面

(逆定理: 严格超平面分离 不成立)

② 支撑超平面: 若 $x_0\in\text{bd}\,C$: $a\ne0$, 且 $\forall x\in C, a^Tx\le a^Tx_0$, 则 $a^Tx=a^Tx_0$ 为支撑超平面

对任意凸集C, $\forall x\in\text{bd}\,C$, 均存在支撑超平面

正定矩阵 $X,Y\succ0$, $\exists Q$, $QXQ^T=\Lambda_X$, $QYQ^T=\Lambda_Y$

① 矩阵凸函数 ② 矩阵凸: $\forall v, v^Tf(A)v$ 凸 ③ 上境图 $(A\in S^n_{++}, v^Tf(A)v\le t)$ 凸 (1) Schur补 证明

函数

凸函数: $f: R^n\to R$, dom f为凸集, $\forall x,y\in\text{dom}\,f$, $0\le\theta\le1$, $f(\theta x+(1-\theta)y)\le\theta f(x)+(1-\theta)f(y)$

Jensen不等式: $f(\theta_1x_1+\dots+\theta_kx_k)\le\theta_1f(x_1)+\dots+\theta_kf(x_k)$

严格凸: $0<\theta<1$, $x\ne y$ 严格不等

二阶条件: dom f凸, $\nabla^2f(x)\succeq0$

一阶条件: f凸 ⇔ $\forall x,y\in\text{dom}\,f$, $f(y)\ge f(x)+\nabla f(x)^T(y-x)$, dom f凸

下水平集: 函数f的α下水平集 $C_\alpha=\{x\in\text{dom}\,f\,|\,f(x)\le\alpha\}$, 对 $\forall\alpha$, 若f凸, 则 C_α 凸

上境图: $\text{epi}\,f=\{(x,t)\,|\,x\in\text{dom}\,f, f(x)\le t\}$; f凸函数 ⇔ 上境图凸集

2. 凸函数例子

e^{ax}, x^α, $\log x$, 范数, $\max(x_1,\dots,x_n)$, $x\log x$, $f(x,y)=x^2/y$, $\log\det X$

3. 保凸运算 (函数)

① 非负加权求和 $f=w_1f_1+\dots+w_nf_n$

② 复合仿射映射 $g(x)=f(Ax+b)$, f凸 ⇒ g凸

③ 逐点最大和逐点上确界: $f_1,\dots,f_m$凸, $f(x)=\max(f_1(x),\dots,f_m(x))$凸 (上境图交集)

④ $f=h(g(x))$: $f(x)=\sup\{g(x)\,|\,\dots\}$
- h凸, 非减, g凸 ⇒ f凸
- h凸, 非增, g凹 ⇒ f凸
- h凹, 非减, g凹 ⇒ f凹
- h凹, 非增, g凸 ⇒ f凹

⑤ 最小化: f(x,y)凸, $g(x)=\inf_{y\in C}f(x,y)$凸 (若 $g(x)>-\infty$)

⑥ 透视函数 $g(x,t)=tf(x/t)$

Schur补: $X=\begin{bmatrix}A&B\\B^T&C\end{bmatrix}$, $S=C-B^TA^{-1}B$

① 广义不等式 (锥偏序): 若K为正常锥 (凸集, 闭集, 实(有非空内部), 尖(不包含直线, $x\in K, -x\in K\Rightarrow x=0$)) ⇒ 广义不等式: $x\preceq_K y\Leftrightarrow y-x\in K$, $x\prec_K y\Leftrightarrow y-x\in\text{int}\,K$

性质: 对加法保序 $x\preceq_K y, u\preceq_K v \Rightarrow x+u\preceq_K y+v$; 传递性 $x\preceq_K y, y\preceq_K z\Rightarrow x\preceq_K z$; 非负数乘 $x\preceq_K y, \alpha\ge0, \alpha x\preceq_K\alpha y$; 反对称 $x\preceq_K y, y\preceq_K x\Rightarrow x=y$

最小元: $\forall y\in S$, $x\preceq_K y$; 极小元: $y\in S, y\preceq_K x \Rightarrow y=x$

② 对偶锥, 广义不等式对偶: $K^*=\{y\,|\,x^Ty\ge0, \forall x\in K\}$

$x\preceq_K y\Leftrightarrow\forall\lambda\succeq_{K^*}0, \lambda^Tx\le\lambda^Ty$

$x\prec_K y\Leftrightarrow\forall\lambda\succeq_{K^*}0, \lambda\ne0, \lambda^Tx<\lambda^Ty$

① K^* 闭凸锥 ② $K_1\subseteq K_2\Rightarrow K_2^*\subseteq K_1^*$ ③ 若K有非空内部, 则 K^* 尖 ④ 若K闭包尖, 则 K^* 有非空内部 ⑤ K^{**} 为K凸包的闭包 (K凸, 闭, $K^{**}=K$)

③ K-凸性: 若 $\forall x,y$, $0\le\theta\le1$, dom f凸, $f(\theta x+(1-\theta)y)\preceq_K\theta f(x)+(1-\theta)f(y)$, K-凸

对偶: f K-凸 ⇔ $\forall w\succeq_{K^*}0$, w^Tf 凸

一阶凸性: $\forall x,y\in\text{dom}\,f$, $f(y)\succeq_K f(x)+Df(x)(y-x)$ ⇔ 凸

复合: $f(x)=h(g(x))$: $g: R^n\to R^p$ K-凸, $h: R^p\to R$ 凸, K-非减 ⇒ $h\circ g$ 凸

1. 问题: minimize $f(x)$ s.t. $f_i(x) \le 0$, $h_i(x) = 0$

要求: $D = \bigcap \mathrm{dom} f_i \cap \mathrm{dom} h_i$ 凸集；$f_0, f_i(x)$ 凸函数；$h_i(x)$ 仿射

凸优化问题

等价变换: 变量代换 $x \to \phi(x)$ (一一映射)；松弛变量: $f_i(x) \le 0 \Leftrightarrow \begin{cases} f_i(x) + s_i = 0 \\ s_i \ge 0 \end{cases}$

最优性条件: 一般: $\nabla f_0(x)^T(y-x) \ge 0, \forall y$；无约束: $\nabla f_0(x) = 0$

2. 常见问题 ① 线性规划 LP: min $c^Tx + d$ s.t. $Gx \le h$, $Ax = b$ (目标、约束仿射)

线性分式规划 转化为线性规划: $y = \frac{x}{e^Tx+f}$, $z = \frac{1}{e^Tx+f}$

② 二次规划 QP: min $\frac{1}{2}x^TPx + q^Tx + r$ s.t. $Gx \le h$, $Ax = b$ (目标二次，约束仿射)

③ 二次约束二次规划 QCQP: min $\frac{1}{2}x^TPx + q^Tx + r$ s.t. $\frac{1}{2}x^TP_ix + q_i^Tx + r_i \le 0$, $Ax = b$ (目标、约束均二次)

④ 二阶锥规划 SOCP: min f^Tx s.t. $\|A_ix + b_i\|_2 \le c_i^Tx + d_i$, $Fx = g$ (C_i 为 0 时, QCQP)

LP ⊂ QP ⊂ QCQP ⊂ SOCP ⊂ SDP ⊂ conic ⊂ 凸优化

⑤ 半定规划 SDP: 不等式 min c^Tx s.t. $x_1F_1 + \dots + x_nF_n + G \preceq B$；标准 min $\mathrm{tr}(CX)$ $\mathrm{tr}(A_iX) = b_i$ $X \succeq 0$ (其中 $X \in S^n$)

3. 几何规划 minimize $f_0(x)$ s.t. $f_i(x) \le 1$, $h_i(x) = 1$

$f_i(x)$ 为正项式 $\sum c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$；$h_i(x)$ 为单项式 $c x_1^{a_1} \cdots x_n^{a_n}$ ($c > 0$, $a_i \in R$)

非凸 → 引入变量, 化为凸: 用 $y_i = \log x_i$ 代替

minimize $\log \sum_{k=1}^{K} e^{a_{0k}^Ty + b_{0k}}$ s.t. $\log \sum e^{a_{ik}^Ty + b_{ik}} \le 0$, $g_i^Ty + h_i = 0$

4. 向量优化 ↓ 多目标 minimize (关于 K 广义不等式) $f_0(x)$ s.t. $f_i(x) \le 0$, $h_i(x) = 0$

① 最优点 = 有最优 唯一最优
② 有效点: 仅有有效点 均为 Pareto 最优值

多目标优化 minimize $F_0(x)$ s.t. $f_i(x) \le 0$, $h_i(x) = 0$ — 有最优点: 唯一；不存在: 求有效；标量化: $\sum \lambda_j F_j(x)$ 最小

标量化: minimize $\lambda^T f_0(x)$ ($\lambda \succ 0$) s.t. $f_i(x) \le 0$, $h_i(x) = 0$

不同的方向 λ, 对应不同的有效点

若为凸问题 每个 Pareto 最优值都可以得到

1. Lagrange 对偶: min $f(x)$ s.t. $f_i(x) \le 0$, $h_i(x) = 0$ ⇒ $L(x,\lambda,\nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum \nu_i h_i(x) \le f_0(x)$, $\lambda_i \ge 0$

⇓ max $g(\lambda,\nu) = \inf_x L(x,\lambda,\nu)$ s.t. $\lambda \ge 0$ 对偶问题

凸优化强对偶 ⇒ 若 Slater 成立 ($\exists x \in \mathrm{relint}\, D$, $f_i(x) < 0$, $Ax = b$) 凸问题 强对偶成立 $p^* = d^*$

⇓ 弱对偶 对偶问题的最优值 $d^* \le p^*$ (原问题最优值)

凸问题 不论原问题是否凸 对偶问题都是凸的

推论: $\sup_{\lambda \ge 0} \inf_x L(x,\lambda) \le \inf_x \sup_{\lambda \ge 0} L(x,\lambda)$ (相当相等 强对偶)

择一定理: 弱择一: 至多一个可行；强择一: 有且仅一个可行

Farkas 引理 强择一: $Ax \le 0, c^Tx < 0$；$A^Ty + c = 0, y \ge 0$

KKT 条件: $\nabla f(x^*) + \sum \lambda_i^* \nabla f_i(x^*) + \sum \nu_i^* \nabla h_i(x^*) = 0$；$f_i(x^*) \le 0$；$h_i(x^*) = 0$；$\lambda_i^* \ge 0$；$\lambda_i^* f_i(x^*) = 0$ 互补松弛

对凸问题 强对偶: 强对偶成立 KKT 为充要条件

凸 + Slater: 强对偶 ⇔ KKT ⇒ 对偶问题最优 $p^* = d^*$

扰动: ⇒ minimize $f_0(x)$ s.t. $f_i(x) \le u_i$ 最优值为 $p^*(u,\nu)$, $h_i(x) = v_i$

有 $p^*(u,v) \ge p^*(0,0) - \lambda^{*T}u - \nu^{*T}v$

$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i}$, $\nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}$ 影子价格

1. 强凸 无约束优化: $\exists m > 0$, 使 $\nabla^2 f(x) \succeq mI$

$f(y) = f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}(y-x)^T\nabla^2 f(z)(y-x) \ge f(x) + \nabla f(x)^T(y-x) + \frac{m}{2}\|y-x\|_2^2$

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

Lipchitz: $\|\nabla f(x) - \nabla f(y)\| \le M\|x-y\| \Leftrightarrow \nabla^2 f(x) \preceq MI$

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

2. 梯度下降: $x = x - t\nabla x$, $\Delta x = -\nabla x$

步长: ① 定步长: $t = \frac{2}{m+M}$ 最优
② 精确直线搜索: $(x + s\Delta x)$, $t = \arg\min_{s \ge 0} f(x + s\Delta x)$
③ 回溯直线搜索: 给定 $\alpha \in (0, 0.5)$, $\beta \in (0,1)$, $t_0 = 1$

~~若 ...~~ 令 $t = \beta t$

若 $f(x + t\Delta x) > f(x) + \alpha t \nabla f(x)^T\Delta x$ 令 $t = \beta t$

收敛性分析: 误差 $f(x^{(k)}) - p^*$ 与 $f(x^{(k-1)}) - p^*$

3. Newton 法: $\Delta x_{nt} = -\nabla^2 f(x)^{-1}\nabla f(x)$

牛顿减量: $\lambda(x) = (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{1/2}$

$\nabla f(x)^T \Delta x_{nt} = -\lambda(x)^2$

Newton 法: 1. Δx_{nt}, $\lambda^2(x)$ 计算 2. 当 $\frac{\lambda^2}{2} \le \epsilon$ 退出 3. 回溯直线搜索确定 t 4. 改进 $x = x + t\Delta x_{nt}$

4. 收敛性: 共两阶段 阻尼 Newton: $f(x^{(k)}) - p^*$ 二次 较慢；二次收敛: $\|\nabla f(x)\|_2 < \eta$ 时

有约束优化 ① 凸二次规划: minimize $\frac{1}{2}x^TPx + q^Tx + r$ s.t. $Ax = b$

KKT 条件: $Ax^* = b$, $Px^* + q + A^T\nu^* = 0$ ⇒ $\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$ KKT 矩阵

② 转为无约束问题: $\{x \mid Ax = b\}$ ⇓ $x = Fz + \hat{x}$ (F 为 $Ax = 0$ 的基, $\hat{x}$ 为 $Ax = b$ 任一特解)

③ 对偶问题: maximize $-b^T\nu - f^*(-A^T\nu)$

④ Newton 法 (可行初始点): Δx_{nt}: $\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \Delta x_{nt} \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$

⑤ Newton 法 (不可行初始点): $\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix} = \begin{bmatrix} -\nabla f(x) - A^T\nu \\ b - Ax \end{bmatrix}$ (使 $A(x + \Delta x) = b$)

$r_{dual} = \nabla f(x) + A^T\nu$, $r_{pri} = Ax - b$, $r(x,\nu) = (r_{dual}, r_{pri})$

不可行初始点 Newton: ① 计算 $\Delta x_{nt}, \Delta \nu_{nt}$ ② 对 $\|r\|_2$ 回溯直线 ③ 改进 x ④ $Ax = b$, $\|r(x,\nu)\|_2 \le \epsilon$ 停止

5. 不等式约束 ① 对数障碍函数 可行 转化为等式约束 min: $f_0(x) + \sum_{i=1}^{m} -\frac{1}{t}\log(-f_i(x))$ s.t. $Ax = b$

$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x))$；minimize $tf_0(x) + \phi(x)$ s.t. $Ax = b$ 对偶问题

对不同 t, 有最优点 $x^*(t)$ 为中心路径, $x^*(t)$ 均严格可行；有 $f_0(x^*(t)) - p^* \le \frac{m}{t}$ (对偶间隙规定, 每次优化只有一个对偶间隙)

② 障碍方法: 给定 min $tf_0 + \phi$ s.t. $Ax = b$ 1. 中心点计算 2. $x = x^*(t)$ 3. 若 $m/t < \epsilon$ 退出 4. 增加 t, $t = \mu t$

原对偶内点法 ① 残差: $r(x,\lambda,\nu) = \begin{bmatrix} \nabla f_0(x) + Df(x)^T\lambda + A^T\nu \\ -\mathrm{diag}(\lambda)f(x) - (1/t)\mathbf{1} \\ Ax - b \end{bmatrix} = \begin{bmatrix} r_{dual} \\ r_{cent} \\ r_{pri} \end{bmatrix}$

$\begin{bmatrix} \nabla^2 f_0(x) + \sum \lambda_i \nabla^2 f_i(x) & Df(x)^T & A^T \\ -\mathrm{diag}(\lambda)Df(x) & -\mathrm{diag}(f(x)) & 0 \\ A & 0 & 0 \end{bmatrix}\begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \nu \end{bmatrix} = -\begin{bmatrix} r_{dual} \\ r_{cent} \\ r_{pri} \end{bmatrix}$

两阶段方法: ① minimize s s.t. $f_i(x) \le s$, $Ax = b$, $s \ge 0$ ② minimize $tf_0(x) - \sum \log(s - f_i(x))$