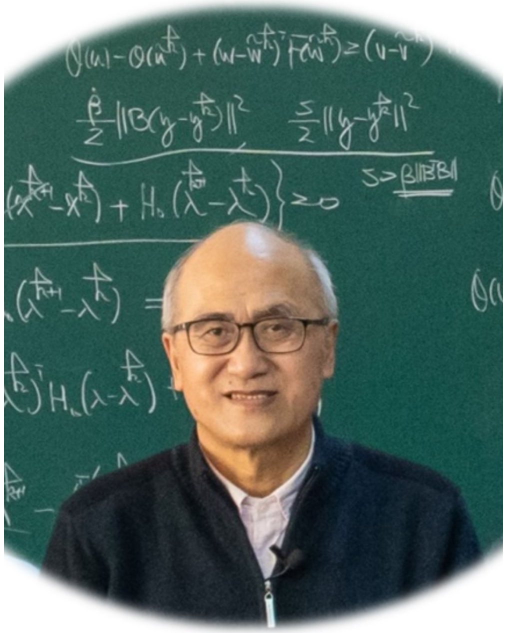
# 凸优化的一些典型问题及其求解方法

## 10. 从变分不等式的PC方法到凸优化的SC方法



南京大学数学系

何炳生

Bingsheng He

http://maths.nju.edu.cn/~hebma/

南京师范大学数学科学学院 2022 年元月 4日 - 9日

从变分不等式(VI) 的投影收缩算法

到凸优化的分裂收缩算法.

都是预测-校正方法

# 1 单调变分不等式

设 $\Omega \subset \Re^n$ 是一个闭凸集, $F$ 是从 $\Re^n$ 到自身的一个算子, 我们讨论单调变分不等式问题

$$\text{VI}(\Omega, F) \qquad u^* \in \Omega, \quad (u - u^*)^T F(u^*) \geq 0, \quad \forall u \in \Omega \tag{1.1}$$

的求解方法. 一个变分不等式称为单调的, 是指 VI$(\Omega, F)$ 中的 $F$ 是 $\Re^n$ (或 $\Omega$) 上的单调算子 (monotone operator), 即 $F$ 满足

$$(u - v)^T (F(u) - F(v)) \geq 0, \quad \forall\, u,\, v \in \Re^n (\text{or } \Omega).$$

说 $F$ 是单调仿射算子, 指 $F(u) = Mu + q,\ q \in \Re^n$, 其中 $M$ 是半正定矩阵. 一个 $n \times n$ 矩阵 $M$ 是半正定的, 是指对任何的 $u \in \Re^n$ 都有

$$u^T M u \geq 0.$$

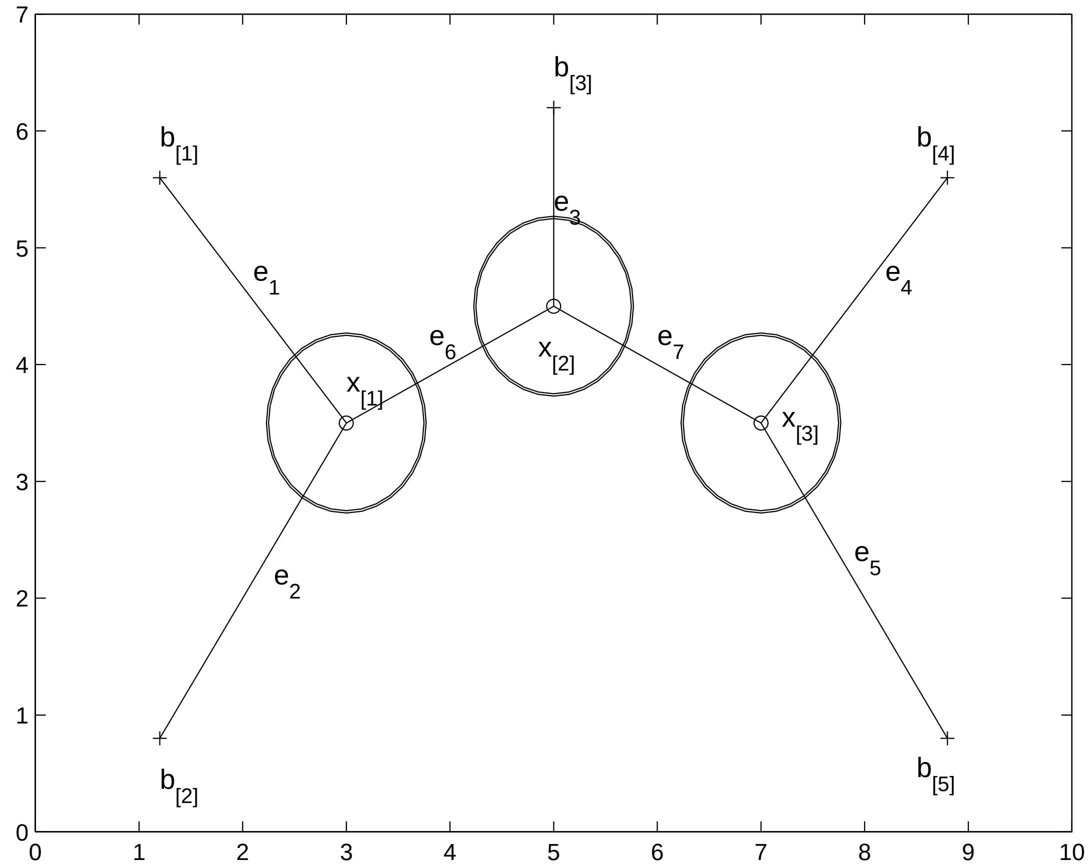
这里并不要求矩阵 $M$ 对称. 换句话说, 只要 $M^T + M$ 对称半正定.

特别是, 当 $M$ 是反对称矩阵, 即 $M^T = -M$ 时, 总有 $u^T M u \equiv 0$. 这时, 仿射算子 $F(u) = Mu + q$ 是单调的.

# 2 线性单调不等式

有些典型的非光滑凸优化问题可以化成结构相当简单的变分不等式. 这一节我们以最短距离和问题为例加以说明.

假设 $b_{[1]}, \ldots, b_{[5]}$ 是确定的村镇. 现在要用一个如右图的网络把它们连接起来. $x_{[1]}, x_{[2]}, x_{[3]}$ 是待选的连接点, 它们必须分别在划定的范围 $X_1, X_2, X_3$ 内. 如何确定 $x_{[1]}, x_{[2]}, x_{[3]}$ 的位置, 使得网络线路长度最短.



我们在 $p$-模意义下求上述网络的最短距离. 该问题的数学模型是

$$\min_{x_{[j]}\in X_j} \left\{ \begin{array}{lll} \|x_{[1]}-b_{[1]}\|_p & +\|x_{[1]}-b_{[2]}\|_p & +\|x_{[2]}-b_{[3]}\|_p \\ & +\|x_{[3]}-b_{[4]}\|_p & +\|x_{[3]}-b_{[5]}\|_p \\ & +\|x_{[1]}-x_{[2]}\|_p & +\|x_{[2]}-x_{[3]}\|_p \end{array} \right\}. \tag{2.1}$$

主要对 $p=1,2,\infty$ 感兴趣. 注意到这里是距离和问题, 不是距离的平方和问题. 此类问题是一个非光滑凸优化问题.

## 2.1 欧氏模下的最短距离和问题

**将问题转化为 min-max 问题.** 注意到对任意的 $d\in\Re^2$, 有

$$\|d\|_2 = \max_{\xi\in B_2} \xi^T d, \tag{2.2}$$

其中

$$B_2 = \{\xi\in\Re^2 \mid \|\xi\|_2\le 1\}.$$

## 利用 (2.2), 上述欧氏模意义下的最短距离和问题可以化为 min-max 问题

$$\min_{x_{[i]}\in X_i}\ \max_{z_{[j]}\in B_2}\left\{\begin{array}{r}z_{[1]}^T(x_{[1]}-b_{[1]})+z_{[2]}^T(x_{[1]}-b_{[2]})+\ z_{[3]}^T(x_{[2]}-b_{[3]})\\ +\ z_{[4]}^T(x_{[3]}-b_{[4]})+\ z_{[5]}^T(x_{[3]}-b_{[5]})\\ +\ z_{[6]}^T(x_{[1]}-x_{[2]})+z_{[7]}^T(x_{[2]}-x_{[3]})\end{array}\right\}$$

它的紧凑形式为

$$\boxed{\min_{x\in\mathcal{X}}\ \max_{z\in\mathcal{B}_2}\ \ z^T(Ax-b)}$$

其中

$$\mathcal{X}=X_1\times X_2\times X_3\ ,\qquad \mathcal{B}_2=B_2\times B_2\times\cdots\times B_2.$$

$$x=\begin{pmatrix}x_{[1]}\\ x_{[2]}\\ x_{[3]}\end{pmatrix},\quad z=\begin{pmatrix}z_{[1]}\\ z_{[2]}\\ \vdots\\ z_{[7]}\end{pmatrix}. \tag{2.3}$$

分块矩阵 $A$ 和向量 $b$ 的结构分别是

$$A = \begin{pmatrix} I_2 & 0 & 0 \\ I_2 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_2 \\ 0 & 0 & I_2 \\ I_2 & -I_2 & 0 \\ 0 & I_2 & -I_2 \end{pmatrix}, \qquad b = \begin{pmatrix} b_{[1]} \\ b_{[2]} \\ b_{[3]} \\ b_{[4]} \\ b_{[5]} \\ 0 \\ 0 \end{pmatrix}.$$

在这个 min-max 问题中, 每个 $x_{[i]}$ 对应一个点, 而每个 $z_{[j]}$ 则对应一条边.

设 $(x^*, z^*) \in \mathcal{X} \times \mathcal{B}_2$ 是 min-max 问题的解, 则对所有的 $x \in \mathcal{X}$ 和 $z \in \mathcal{B}$, 有

$$\boxed{z^T(Ax^* - b) \le {z^*}^T(Ax^* - b) \le {z^*}^T(Ax - b)}$$

它的等价形式是下面的线性变分不等式:

$$x^* \in \mathcal{X},\ z^* \in \mathcal{B}_2, \quad \begin{cases} (x - x^*)^T (A^T z^*) \geq 0, & \forall x \in \mathcal{X}, \\ (z - z^*)^T (-Ax^* + b) \geq 0, & \forall z \in \mathcal{B}_2. \end{cases}$$

可以写成更简单紧凑的形式

$$u^* \in \Omega, \quad (u - u^*)^T (Mu^* + q) \geq 0, \quad \forall u \in \Omega$$

其中有

$$u = \begin{pmatrix} x \\ z \end{pmatrix}, \quad M = \begin{pmatrix} 0 & A^T \\ -A & 0 \end{pmatrix}, \quad q = \begin{pmatrix} 0 \\ b \end{pmatrix}$$

和

$$\Omega = \mathcal{X} \times \mathcal{B}_2.$$

## 2.2 $l_1$-模下的最短距离和问题

由于对任意的 $d \in \Re^2$, 有

$$\|d\|_1 = \max_{\xi \in B_\infty} \xi^T d,$$

其中

$$B_\infty = \{\xi \in R^2 \mid \|\xi\|_\infty \leq 1\}.$$

与欧氏模下的最短距离和问题一样, 问题 (2.1) 在 $l_1$-模意义下也可以表示成一个 min-max 问题

$$\min_{x_{[i]} \in X_i} \max_{z_{[j]} \in B_\infty} \left\{ \begin{array}{r} z_{[1]}^T(x_{[1]} - b_{[1]}) + z_{[2]}^T(x_{[1]} - b_{[2]}) + z_{[3]}^T(x_{[2]} - b_{[3]}) \\ + z_{[4]}^T(x_{[3]} - b_{[4]}) + z_{[5]}^T(x_{[3]} - b_{[5]}) \\ + z_{[6]}^T(x_{[1]} - x_{[2]}) + z_{[7]}^T(x_{[2]} - x_{[3]}) \end{array} \right\}$$

它的紧凑形式为

$$\boxed{\min_{x \in \mathcal{X}} \max_{z \in \mathcal{B}_\infty} z^T(Ax - b)}$$

与欧氏模下的距离和问题一样, 它等价于变分不等式

$$\boxed{u^* \in \Omega, \quad (u-u^*)^T(Mu^*+q) \geq 0, \quad \forall u \in \Omega}$$

矩阵 $M$ 和向量 $q$ 都不变, 所不同的只是这时集合

$$\Omega = \mathcal{X} \times \mathcal{B}_\infty, \qquad \mathcal{B}_\infty = B_\infty \times B_\infty \times \cdots \times B_\infty.$$

## 2.3 $l_\infty$-模下的最短距离和问题

由于对任意的 $d \in \Re^2$, 有

$$\|d\|_\infty = \max_{\xi \in B_1} \xi^T d,$$

其中

$$B_1 = \{\xi \in R^2 \mid \|\xi\|_1 \leq 1\}.$$

与欧氏模下的最短距离和问题一样, 问题 (2.1) 在 $l_\infty$-模意义下也可以表示成

一个 min-max 问题

$$\min_{x_{[i]}\in X_i}\max_{z_{[j]}\in B_1}\left\{\begin{array}{r} z_{[1]}^T(x_{[1]}-b_{[1]})+z_{[2]}^T(x_{[1]}-b_{[2]})+\ z_{[3]}^T(x_{[2]}-b_{[3]}) \\ +\ z_{[4]}^T(x_{[3]}-b_{[4]})+\ z_{[5]}^T(x_{[3]}-b_{[5]}) \\ +\ z_{[6]}^T(x_{[1]}-x_{[2]})+\ z_{[7]}^T(x_{[2]}-x_{[3]}) \end{array}\right\}$$

并化成等价的变分不等式

$$\boxed{u^*\in\Omega,\quad (u-u^*)^T(Mu^*+q)\ge 0,\quad \forall u\in\Omega}$$

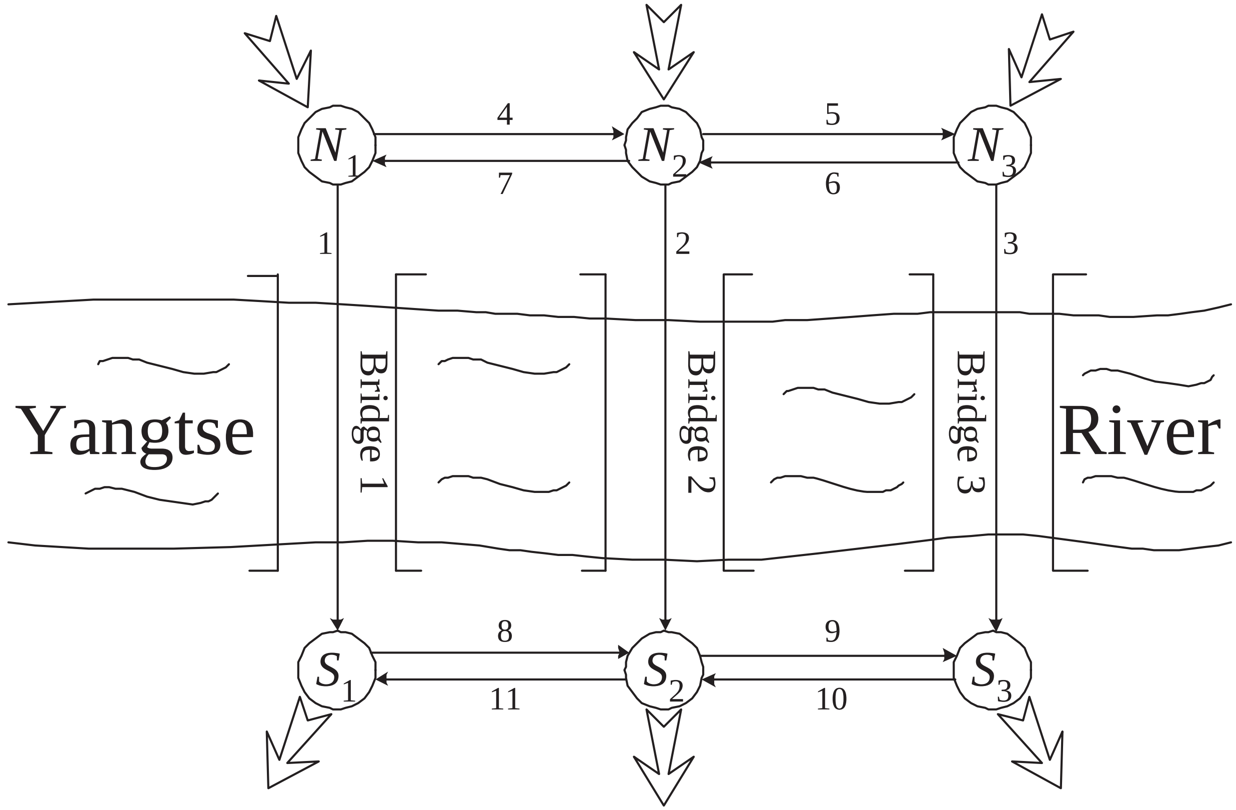
矩阵 $M$ 和向量 $q$ 都不变, 所不同的只是这时集合

$$\Omega=\mathcal{X}\times\mathcal{B}_1,\qquad \mathcal{B}_1=B_1\times B_1\times\cdots\times B_1.$$

将处理欧氏模问题时的 $\mathcal{B}_2$ 换成了$\mathcal{B}_1$.

# 3 交通疏导问题之互补问题

设某跨江城市有三座长江大桥, 分别为 Bridge-1, Bridge-2 和 Bridge-3.




不失一般性, 我们可以将 $N_1, N_2, N_3$ 看作由北向南的车辆在江北的出发地, 把 $S_1, S_2, S_3$ 看作它们在江南的集散地.

我们只对管理部门的问题感兴趣, 假设只收过桥费. 管理部门想制定一个适当的收费标准合理控制桥上流量.

## 驾驶员基于 Wardrop 原理的最优出行方案 — 最小费用路径

对给定的大桥收费 $x=(x_1,x_2,x_3)$, 驾驶员会找到他们的最优出行方案.

## 管理部门要求通过大桥合理收费控制桥上的流量

- $0 \le x \in \Re^3$: 桥上的收费向量;
- $f(x) \in \Re^3$: 桥上的流量, 它是收费 $x$ 的函数;
- $0 < b \in \Re^3$: 管理部门希望控制的桥上的流量上界.

### 管理部门要求解的数学问题是

$$x \ge 0, \qquad F(x) = b - f(x) \ge 0, \qquad x^T F(x) = 0.$$

同样, 流量 $f(x)$ 确是收费 $x$ 的函数, 但没有表达式. 只能对给定的自变量, 观测相应的函数值, 而这种观测, 往往代价不菲. 我们从事投影收缩算法研究, 着眼点是要得到效率高一些的、只用函数值和少用函数值的方法.