# Intorduction to Interior Point Methods

Jakub Kruzik[1,2]
jakub.kruzik@ugn.cas.cz

16 February 2022
Mathematics and Computer Science Seminar, UGN CAS

[1]Institute of Geonics of the Czech Academy of Sciences
[2]Dept. of Appl. Math., Faculty of Electr. Engin. and CS, VSB - TU Ostrava

# Problem definition QP (LP)

$$\underset{\boldsymbol{x}}{\operatorname{argmin}}\, f(\boldsymbol{x}) \quad \text{s.t.} \quad x \in \Omega$$

$$\underset{\boldsymbol{x}}{\operatorname{argmin}}\, \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{c} \quad \text{s.t.} \quad \boldsymbol{x} \geq \boldsymbol{o}, \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{1}$$

where $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ is SPS, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ full row rank $m \leq n$, $\boldsymbol{Q} = \boldsymbol{O} \rightarrow$ LP.

# Problem definition QP (LP)

$$\underset{\boldsymbol{x}}{\operatorname{argmin}}\, f(\boldsymbol{x}) \quad \text{s.t.} \quad x \in \Omega$$

$$\underset{\boldsymbol{x}}{\operatorname{argmin}}\, \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{c} \quad \text{s.t.} \quad \boldsymbol{x} \geq \boldsymbol{o}, \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{1}$$

where $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ is SPS, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ full row rank $m \leq n$, $\boldsymbol{Q} = \boldsymbol{O} \rightarrow$ LP.

Dual:

$$\underset{(\boldsymbol{y},\boldsymbol{s})}{\operatorname{argmax}} -\frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{y}^T\boldsymbol{b} \quad \text{s.t.} \quad \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} = \boldsymbol{c}, \quad \boldsymbol{y} \in \mathbb{R}^m, \quad \boldsymbol{s} \geq \boldsymbol{o}. \tag{1}$$

$\boldsymbol{x}$, $(\boldsymbol{y}, \boldsymbol{s})$ is dual-primal pair of QP/LP.

# First order optimality conditions

Lagrangian:

$$L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{y}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) - \boldsymbol{s}^T\boldsymbol{x}$$

# First order optimality conditions

Lagrangian:

$$L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{y}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) - \boldsymbol{s}^T\boldsymbol{x}$$

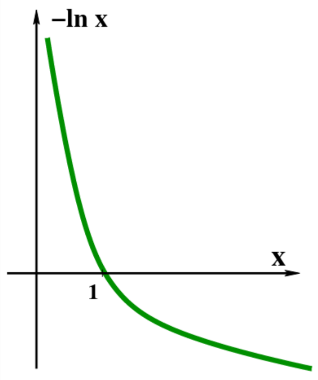
KKT:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b}, \\ \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} &= \boldsymbol{c}, \\ \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} &= \boldsymbol{o}, \\ \boldsymbol{x} &\geq \boldsymbol{o}, \\ \boldsymbol{s} &\geq \boldsymbol{o}, \end{aligned}$$

where $\boldsymbol{X} = diag(\boldsymbol{x})$, $\boldsymbol{S} = diag(\boldsymbol{s})$, $\boldsymbol{e} = (1, 1, \ldots, 1)$.

# Towards IPM (1) - Logarithmic Barrier

Enforce $x_i \geq 0$ using penalty with $-\ln x_i$



# Towards IPM (2) - Logarithmic Barrier Formulation

$$\underset{\boldsymbol{x}}{\operatorname{argmin}} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{c} - \mu\sum_{i=1}^{n}\ln \boldsymbol{x}_i \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{2}$$

$$\underset{(\boldsymbol{y},\boldsymbol{s})}{\operatorname{argmax}} -\frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{y}^T\boldsymbol{b} + \mu\sum_{i=1}^{n}\ln \boldsymbol{s}_i \quad \text{s.t.} \quad \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} = \boldsymbol{c}, \quad \boldsymbol{y} \in \mathbb{R}^m. \tag{2}$$

# Towards IPM (2) - Logarithmic Barrier Formulation

$$\underset{\boldsymbol{x}}{\operatorname{argmin}} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{c} - \mu\sum_{i=1}^{n}\ln \boldsymbol{x}_i \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{2}$$

$$\underset{(\boldsymbol{y},\boldsymbol{s})}{\operatorname{argmax}} -\frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{y}^T\boldsymbol{b} + \mu\sum_{i=1}^{n}\ln \boldsymbol{s}_i \quad \text{s.t.} \quad \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} = \boldsymbol{c}, \quad \boldsymbol{y} \in \mathbb{R}^m. \tag{2}$$

$$L(\boldsymbol{x}, \boldsymbol{y}, \mu) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{y}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) - \mu\sum_{i=1}^{n}\ln \boldsymbol{x}_i$$

# Towards IPM (2) - Logarithmic Barrier Formulation

$$\underset{\boldsymbol{x}}{\text{argmin}}\ \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{c} - \mu\sum_{i=1}^{n}\ln \boldsymbol{x}_i \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{2}$$

$$\underset{(\boldsymbol{y},\boldsymbol{s})}{\text{argmax}} -\frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{y}^T\boldsymbol{b} + \mu\sum_{i=1}^{n}\ln \boldsymbol{s}_i \quad \text{s.t.} \quad \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} = \boldsymbol{c}, \quad \boldsymbol{y} \in \mathbb{R}^m. \tag{2}$$

$$L(\boldsymbol{x},\boldsymbol{y},\mu) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{y}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) - \mu\sum_{i=1}^{n}\ln \boldsymbol{x}_i$$

KKT:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b}, \\ \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} &= \boldsymbol{c}, \\ \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} &= \mu\boldsymbol{e}, \\ \boldsymbol{x} \geq \boldsymbol{o}, \boldsymbol{s} &\geq \boldsymbol{o}. \end{aligned}$$

# Towards IPM (3) - Solving KKT System

KKT:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b}, \\ \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} &= \boldsymbol{c}, \\ \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} &= \mu\boldsymbol{e}, \\ \boldsymbol{x} &\geq \boldsymbol{o}, \boldsymbol{s} \geq \boldsymbol{o}. \end{aligned}$$

As $\mu \to 0$ formulation (2) converges to the solution of (1)

# Towards IPM (3) - Solving KKT System

KKT:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b}, \\ \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} &= \boldsymbol{c}, \\ \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} &= \mu\boldsymbol{e}, \\ \boldsymbol{x} \geq \boldsymbol{o}&, \boldsymbol{s} \geq \boldsymbol{o}. \end{aligned}$$

As $\mu \to 0$ formulation (2) converges to the solution of (1)

Single step of Newton:

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} & \boldsymbol{O} \\ -\boldsymbol{Q} & \boldsymbol{A}^T & \boldsymbol{I} \\ \boldsymbol{S} & \boldsymbol{O} & \boldsymbol{X} \end{pmatrix} \begin{pmatrix} \Delta\boldsymbol{x} \\ \Delta\boldsymbol{y} \\ \Delta\boldsymbol{s} \end{pmatrix} = \begin{pmatrix} \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} \\ \boldsymbol{x} - \boldsymbol{Q}\boldsymbol{x} - \boldsymbol{A}^T\boldsymbol{y} - \boldsymbol{s} \\ \sigma\mu\boldsymbol{e} - \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} \end{pmatrix} = \begin{pmatrix} \xi_p \\ \xi_d \\ \xi_\mu \end{pmatrix} \tag{3}$$

where $\sigma \in (0, 1)$ is barrier reduction parameter.

# Infeasible Path-following Primal-dual IPM - Initialization

**Parameters**

$\alpha_0 = 0.99$ a fraction-to-the-boundary stepsize factor;
$\sigma \in (0,1)$ barrier reduction parameter;
$\varepsilon_p$, $\varepsilon_d$, $\varepsilon_o$ primal feasibility, dual feasibility and optimality tolerances:
IPM stops when $\frac{\|\xi_p^k\|}{1+\|b\|} \leqslant \varepsilon_p$, $\frac{\|\xi_d^k\|}{1+\|c\|} \leqslant \varepsilon_d$ and
$\frac{(x^k)^T s^k/n}{1+|c^T x^k + 1/2 (x^k)^T Q x^k|} \leqslant \varepsilon_o$.

**Initialize IPM**

iteration counter $k = 0$;
primal–dual point $x^0 > 0$, $y^0 = 0$, $s^0 > 0$;
barrier parameter $\mu^0 = (x^0)^T s^0/n$;
primal and dual infeasibilities $\xi_p^0 = b - Ax^0$ and
$\xi_d^0 = c - A^T y^0 - s^0 + Qx^0$.

**Interior Point Method**

**while** $\left(\frac{\|\xi_p^k\|}{1+\|b\|} > \varepsilon_p \text{ or } \frac{\|\xi_d^k\|}{1+\|c\|} > \varepsilon_d \text{ or } \frac{(x^k)^T s^k/n}{1+|c^T x^k + 1/2 (x^k)^T Q x^k|} > \varepsilon_o\right)$ **do**

Update (reduce) the barrier $\mu^{k+1} = \sigma\mu^k$;

Solve the KKT system (7): find the primal–dual Newton direction $(\Delta x, \Delta y, \Delta s)$.

Find $\alpha_P = \max\{\alpha : x^k + \alpha\Delta x \geqslant 0\}$ and $\alpha_D = \max\{\alpha : s^k + \alpha\Delta s \geqslant 0\}$;

Set $\alpha_P := \alpha_0\alpha_P$ and $\alpha_D := \alpha_0\alpha_D$;

Make step

$x^{k+1} = x^k + \alpha_P\Delta x$;

$y^{k+1} = y^k + \alpha_D\Delta y$;

$s^{k+1} = s^k + \alpha_D\Delta s$.

Compute the infeasibilities: $\xi_p^{k+1} = b - Ax^{k+1}$ and $\xi_d^{k+1} = c - A^T y^{k+1} - s^{k+1} + Qx^{k+1}$;

Update the iteration counter: $k := k + 1$.

**end-while**

# Centrality Meassure and Complementary Product - $\mu$

Primal dual iterates of feasible IPM belong to

$$\mathcal{F}^0 = \left\{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) \,|\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{Q}\boldsymbol{x} = c, (\boldsymbol{x}, \boldsymbol{s}) > \boldsymbol{o}\right\},$$

If iterates are kept in neighbourhood of the central path:

$$\mathrm{N}_2(\theta) = \left\{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) \in \mathcal{F}^0 | \|\boldsymbol{X}\boldsymbol{S}\boldsymbol{e} - \mu\boldsymbol{e}\| \leq \theta\mu\right\},$$

then $\varepsilon$-accurate solution is achieved in $\mathcal{O}\left(\sqrt{n}\ln(1/\varepsilon)\right)$.

# Feasible VS Infeasible IPM

In infeasible IPM the Newton direction in (3) is computed with

$$\begin{pmatrix} \xi_p \\ \xi_d \\ \xi_\mu \end{pmatrix},$$

while feasible IPM uses only

$$\begin{pmatrix} \boldsymbol{o} \\ \boldsymbol{o} \\ \xi_\mu \end{pmatrix}$$

as the right-hand side.

# Methora Predictor-corrector

Ignore centrality by setting $\sigma = 0$ in (3) and compute $\Delta^{pred}$. I.e., new RHS is

$$\begin{pmatrix} \xi_p \\ \xi_d \\ -\boldsymbol{XSe} \end{pmatrix}.$$

If full step in $\Delta^{pred}$ is taken, the complementarity product would be

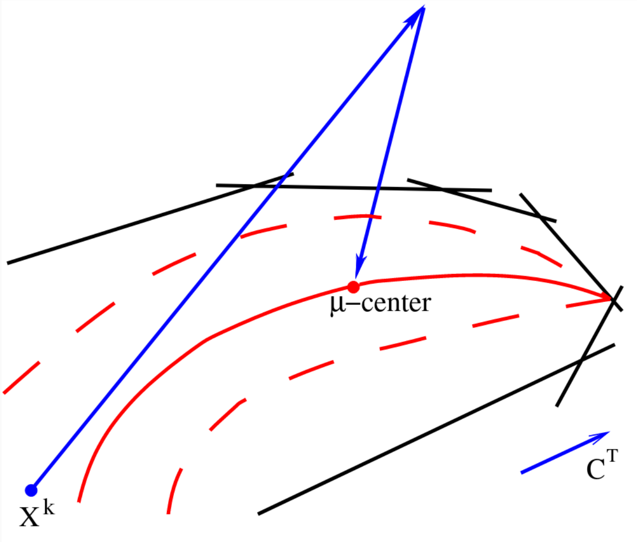
$$(\boldsymbol{X} + \Delta\boldsymbol{X})(\boldsymbol{S} + \Delta\boldsymbol{S})\boldsymbol{e} = \cdots = \Delta\boldsymbol{X}\Delta\boldsymbol{S}\boldsymbol{e} \neq \sigma\mu\boldsymbol{e}.$$

To correct this, we compute $\Delta^{corr}$ solving (3) with RHS

$$\begin{pmatrix} \boldsymbol{o} \\ \boldsymbol{o} \\ \sigma\mu\boldsymbol{e} - \Delta\boldsymbol{X}\Delta\boldsymbol{S}\boldsymbol{e} \end{pmatrix}.$$

Do Newton step in $\Delta = \Delta^{pred} + \Delta^{corr}$.

# Methora Predictor-corrector Illustration



# Solving Linear System in Newton

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} & \boldsymbol{O} \\ -\boldsymbol{Q} & \boldsymbol{A}^T & \boldsymbol{I} \\ \boldsymbol{S} & \boldsymbol{O} & \boldsymbol{X} \end{pmatrix} \begin{pmatrix} \Delta \boldsymbol{x} \\ \Delta \boldsymbol{y} \\ \Delta \boldsymbol{s} \end{pmatrix} = \begin{pmatrix} \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} \\ \boldsymbol{x} - \boldsymbol{Q}\boldsymbol{x} - \boldsymbol{A}^T\boldsymbol{y} - \boldsymbol{s} \\ \sigma\mu\boldsymbol{e} - \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} \end{pmatrix} = \begin{pmatrix} \xi_p \\ \xi_d \\ \xi_\mu \end{pmatrix}$$

is usually reduced by elliminating

$$\Delta \boldsymbol{s} = \boldsymbol{X}^{-1}(\xi_\mu - \boldsymbol{S}\Delta\boldsymbol{x})$$

into symmetric indefinite system

$$\begin{pmatrix} -\boldsymbol{Q} - \boldsymbol{\Theta}^{-1} & \boldsymbol{A}^T \\ \boldsymbol{A} & \boldsymbol{O} \end{pmatrix} \begin{pmatrix} \Delta \boldsymbol{x} \\ \Delta \boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \xi_d - \boldsymbol{X}^{-1}\xi_\mu \\ \xi_p \end{pmatrix},$$

where $\boldsymbol{\Theta}^{-1} = \boldsymbol{X}\boldsymbol{S}^{-1}$.

# Preliminary Numerical Experiments

# String Displacement with Lower Bound

Finited difference discretization of

$$-u''(x) = -15, \quad x \in [0,1] \qquad s.t. \quad u(x) \geq \frac{sin(4\pi x - \frac{\pi}{6})}{2} - 2,$$

| Method | $n$ | Iterations | Hessian Multiplications |
|---|---:|---:|---:|
| MPRGP | 100 | - | 216 |
| IPM | 100 | 12 | - |
| IPM CG | 100 | 12 | 993 |
| MPRGP | 1,000 | - | 3,205 |
| IPM | 1,000 | 17 | - |
| IPM CG | 1,000 | 17 | 12,142 |
| MPRGP | 5,000 | - | 36,941 |
| IPM | 5,000 | 20 | - |
| IPM CG | 5,000 | 12 | 63,894 |

# IPM Progression $n = 5000$ (1)

| Iter | Infeas | mu | PCG Iter |
|---|---|---|---|
| 1 | 4.16E+03 | 2.47E-02 | 24 |
| 2 | 2.08E+01 | 6.97E-04 | 376 |
| 3 | 1.04E-01 | 1.23E-04 | 2500 |
| 4 | 5.25E-03 | 2.39E-05 | 4897 |
| 5 | 2.61E-03 | 5.36E-06 | 4906 |
| 6 | 1.31E-03 | 1.33E-06 | 4764 |
| 7 | 3.79E+01 | 3.60E-07 | 4528 |
| 8 | 1.89E-01 | 8.94E-08 | 4423 |
| 9 | 1.64E+01 | 2.43E-08 | 4325 |
| 10 | 1.70E+01 | 7.05E-09 | 3539 |

# IPM Progression $n = 5000$ (2)

| | | | |
|---|---|---|---|
| 11 | 7.20E+00 | 2.00E-09 | 3579 |
| 12 | 3.49E+00 | 5.71E-10 | 3514 |
| 13 | 1.71E+00 | 1.65E-10 | 3522 |
| 14 | 7.49E-01 | 4.69E-11 | 3344 |
| 15 | 3.11E-01 | 1.31E-11 | 3285 |
| 16 | 1.50E-01 | 3.64E-12 | 3230 |
| 17 | 5.75E-02 | 9.90E-13 | 2693 |
| 18 | 2.88E-04 | 2.52E-13 | 2507 |
| 19 | 1.56E-06 | 6.36E-14 | 2211 |
| 20 | 7.05E-07 | 1.61E-14 | 1727 |
| | | | 63894 |

# Journal Bearing Problem

| Method | Grid | Iterations | Hessian Multiplications |
|---|---|---|---|
| MPRGP | 50x50 | - | 212 |
| IPM | 50x50 | 16 | - |
| IPM CG | 50x50 | 17 | 1,642 |
| MPRGP | 100x100 | - | 430 |
| IPM CG | 100x100 | 20 | 3,535 |
| MPRGP | 200x50 | - | 824 |
| IPM CG | 200x50 | 22 | 6,918 |
| MPRGP | 400x25 | - | 2,193 |
| IPM CG | 400x25 | 24 | 14,253 |

# Cube Contact Problem with FETI

| Method | Outer Iterations | Inner Iterations |
|---|---:|---:|
| MPRGP | 7 | 75 |
| IPM SMALXE | 11 | 90 |
| IPM | 30 | - |
| IPM SMALXE PC | 10 | 37 |
| IPM PC | 8 | - |

# Conclusion and Outlook

- Very good (polynomial) convergence
- Good numerical scalability (outer iterations)
- Predictor-corrector schemes improve convergence
- **Preconditioning is difficult**

Continued collaboration with Jacek Gondzio's group on IPM for contact problems.

# Thank you for your attention!
# Any questions?

Jakub Kruzik[1,2]
jakub.kruzik@ugn.cas.cz

http://permon.vsb.cz

[1]Institute of Geonics of the Czech Academy of Sciences
[2]Dept. of Appl. Math., Faculty of Electr. Engin. and CS, VSB - TU Ostrava