School of Mathematics

Towards an Interior-Point SQP Scheme with Warmstarts

Andreas Grothey, Jacek Gondzio

Overview

- Where we are
  - OOPS (Parallel Structure Exploiting IPM)
  - Warmstarting (Crash Start, Efficient Frontier, nonlinear ALM)
- Where we want to be
  - Concept of l_2 IPM-SQP: Solving sequence of QPs by IPM
- But this is a bad idea, why proceed?
  - Experience in warmstarting IPMs
  - Global/local convergence properties of SQP-type methods
  - If it wasn't, how could it be done?
- What still needs to be done
  - Robust implementation

Interested in large scale solutions

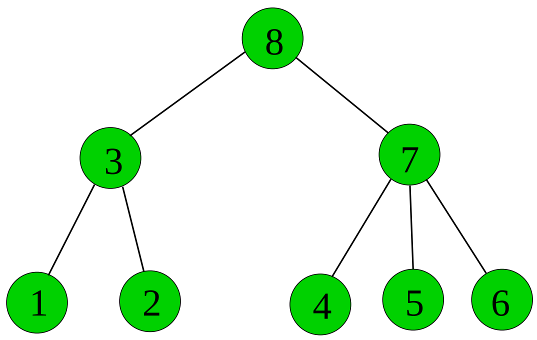
Where we are: OOPS (Object Oriented Parallel Solver)

- Mantra: "Truly large scale problems are not only sparse but structured" (due to e.g. dynamics, uncertainty, spatial distribution etc.)
- Exploiting structure is key to building efficient IPMs for large problems:
  - Faster linear algebra
  - Reduced memory use
  - Possibility to exploit (massive) parallelism
  - **We assume that structure is known!** $\Rightarrow$ no automatic detection.
- OOPS currently solves LP/QP problems.
- Simple sequential-QP scheme solves nonlinear ALM models

OOPS: (Block) Elimination Trees:

Elimination tree orders rows/columns for elimination with minimum fill-in:

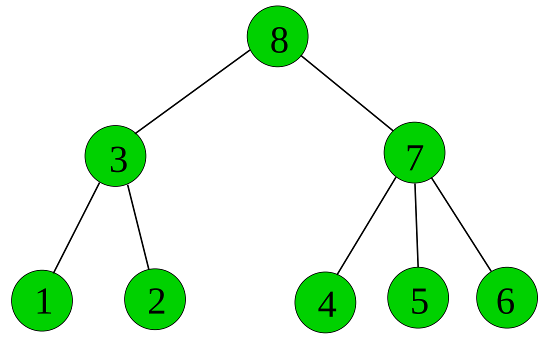
$$\begin{bmatrix} x & & x & & & & & x \\ & x & x & & & & & x \\ x & x & x & & & & & x \\ & & & x & & & x & x \\ & & & & x & & x & x \\ & & & & & x & x & x \\ & & & x & x & x & x & x \\ x & x & x & x & x & x & x & x \end{bmatrix}$$



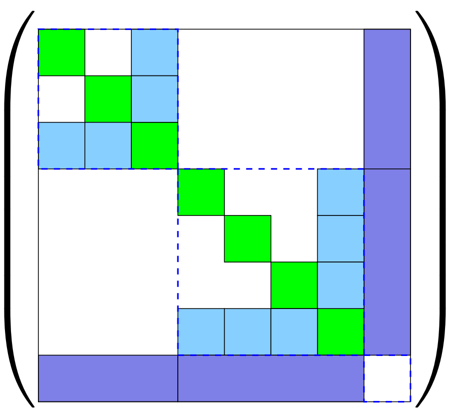
OOPS: (Block) Elimination Trees:

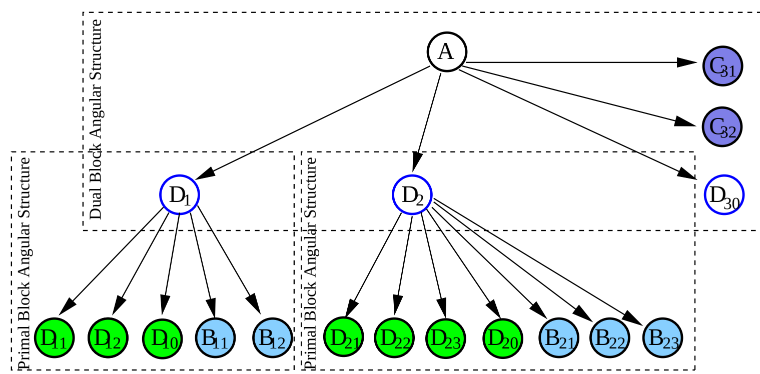
Elimination tree orders rows/columns for elimination with minimum fill-in:

$$\begin{bmatrix} x & & x & & & & & x \\ & x & x & & & & & x \\ x & x & x & & & & & x \\ & & & x & & & x & x \\ & & & & x & & x & x \\ & & & & & x & x & x \\ & & & x & x & x & x & x \\ x & x & x & x & x & x & x & x \end{bmatrix}$$



Elimination Tree can be extended to Block Elimination Tree

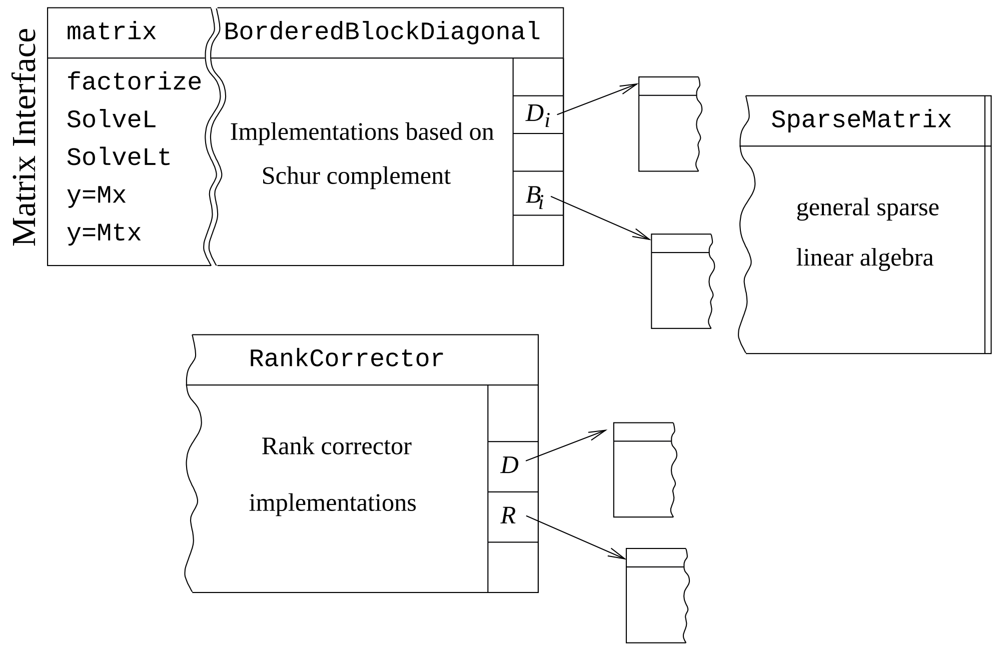




⇒ Organisation of linear algebra, Parallelism

OOPS: Object-oriented linear algebra implementation

- Every node in *block elimination tree* has own linear algebra implementation (depending on its type)
- Implementation is realisation of an abstract linear algebra interface.
- Different implementations for different structures are available.



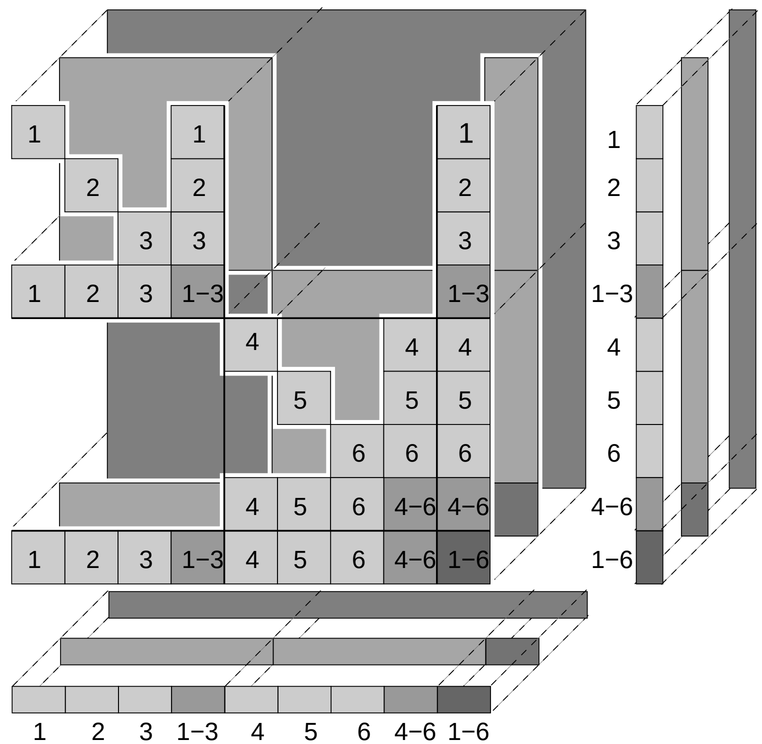
⇒ Rebuild *block elimination tree* with matrix interface structures

Memory Management

Distribution of leading matrix blocks among processors implies

- Distribution of subordinate blocks
- Distribution of row/column vector contributions

⇒ Facilitates exploitation of massive parallelism



Application: Asset and Liability Management Problem - ALM

- A set of assets $\mathcal{J} = \{1, \ldots, J\}$ is given (e.g. bonds, stock, real estate).
- At every stage $t = 0, \ldots, T-1$ we can buy or sell different assets.
- The return of asset j at stage t is *uncertain* (but distribution is known).

We have to make investment decisions: what to buy or sell, at which time stage

Objectives:

- maximize the final wealth
- minimize the associated risk

$\Rightarrow$ Mean Variance formulation:

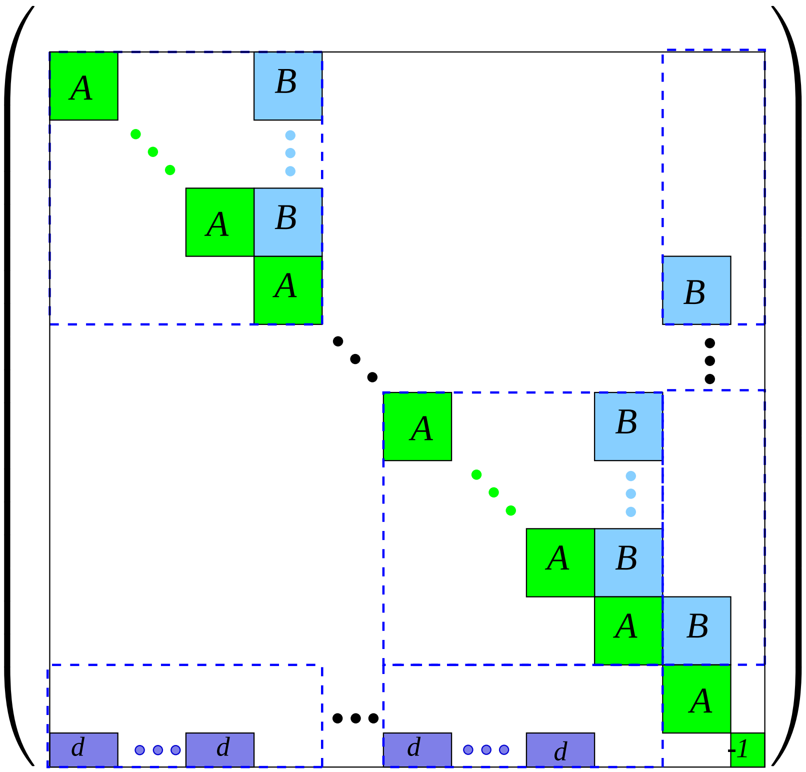
$$\max \mathbb{E}(X) - \rho \mathrm{Var}(X)$$

$\Rightarrow$ Stochastic Program:

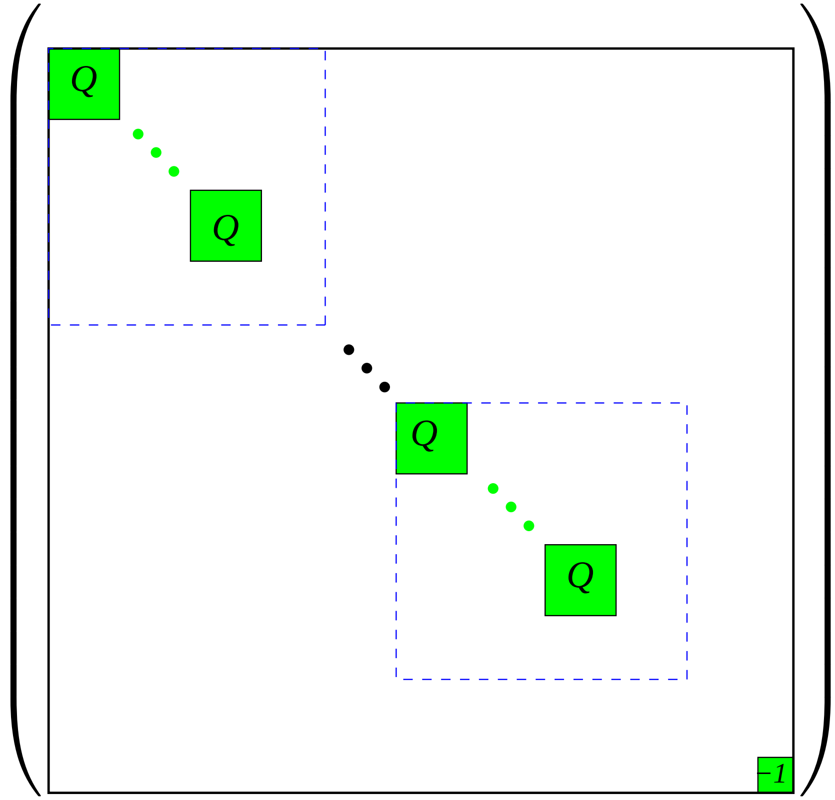
- Can formulate deterministic equivalent problem
- standard QP, but huge

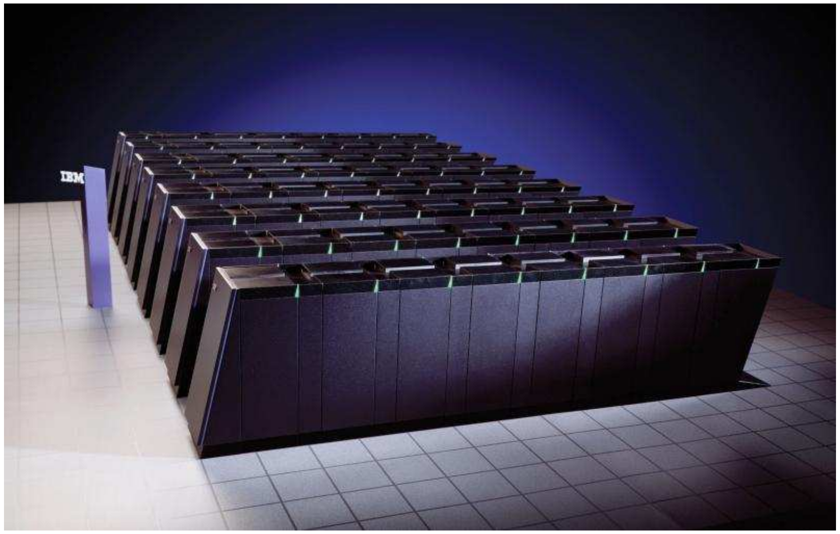
ALM: Structure of matrices A and Q:

Matrix A



Matrix Q





BlueGene (Edinburgh, Scotland)

- 2048 Processors
- 0.7GHz, 256Mb
- 4.7 GFlops
- #33 in `top500.org` list

HPCx (Daresbury, England)

- 1600 IBM Power-4 Processors
- 1.7GHz, 800Mb
- 6.6 GFlops
- #27 in `top500.org` list

Results (ALM: Mean-Variance QP formulation):

Problem	Stages	Blk	Assets	Scenarios	Constraints	Variables	iter	time	procs	machine
ALM1	5	10	5	11.111	66.667	166.666	14	86	1	SunFire 15K
ALM2	6	10	5	111.111	666.667	1.666.666	22	387	5	“
ALM3	6	10	10	111.111	1.222.222	3.333.331	29	1638	5	“
ALM4	5	24	5	346.201	2.077.207	5.193.016	33	856	8	“
UNS1	5	35	5	360.152	2.160.919	5.402.296	27	872	8	“
ALM5	4	64	12	266.305	3.461.966	9.586.981	18	1195	8	“
ALM6	4	120	5	1.742.521	10.455.127	26.137.816	18	1470	16	“
ALM7	4	120	10	1.742.521	19.167.732	52.275.631	19	8465	16	“

Results (ALM: Mean-Variance QP formulation):

Problem	Stages	Blk	Assets	Scenarios	Constraints	Variables	iter	time	procs	machine
ALM1	5	10	5	11.111	66.667	166.666	14	86	1	SunFire 15K
ALM2	6	10	5	111.111	666.667	1.666.666	22	387	5	“
ALM3	6	10	10	111.111	1.222.222	3.333.331	29	1638	5	“
ALM4	5	24	5	346.201	2.077.207	5.193.016	33	856	8	“
UNS1	5	35	5	360.152	2.160.919	5.402.296	27	872	8	“
ALM5	4	64	12	266.305	3.461.966	9.586.981	18	1195	8	“
ALM6	4	120	5	1.742.521	10.455.127	26.137.816	18	1470	16	“
ALM7	4	120	10	1.742.521	19.167.732	52.275.631	19	8465	16	“
ALM8	7	128	6	12.831.873	64.159.366	153.982.477	42	3923	512	BlueGene
ALM9	7	64	14	6.415.937	96.239.056	269.469.355	39	4692	512	BlueGene
ALM10	7	128	13	12.831.873	179.646.223	500.443.048	45	6089	1024	BlueGene
?					?	?				HPCx

Warmstarting (infeasible) Interior Point Methods

$$\begin{array}{rl} \min & c^\top x + \frac{1}{2}x^\top Q x \\ \text{s.t.} & Ax = b \\ & 0 \le x \le u \end{array} \qquad \text{(QP)} \qquad\qquad \begin{array}{rl} \min & \tilde{c}^\top x + \frac{1}{2}x^\top \tilde{Q} x \\ \text{s.t.} & \tilde{A}x = \tilde{b} \\ & 0 \le x \le \tilde{u} \end{array} \qquad (\tilde{\text{Q}}\text{P})$$

Task: From an (approximate) solution of (QP) construct a starting point for $(\tilde{\text{Q}}\text{P})$.

- Aim for centrality and approximate primal/dual feasibility
- the bigger μ and the more balanced $x_i z_i$,
  the more infeasibility can be absorbed in one step

Possible strategies:

- Save approximate μ-center of (QP) and *modify*.
- Only attempt to absorb part of the infeasibility at every iteration

Warmstarting Interior Point Methods II: Modifications

- Can absorb changes in objective into z:

$$\begin{aligned} A^\top y + z - Qx &= c \\ Ax &= b \\ XZe &= \mu e \\ x, z &\geq 0 \end{aligned}$$

  $\Rightarrow$ Changes from c, Q to $\tilde{c}, \tilde{Q}$ can be absorbed into z as long as $z > 0(\sqrt{\mu})$.

- Balancing complementarity products:
  project (x_i, z_i) onto $\{\mu/\sigma \leq x_i z_i \leq \sigma\mu\}$ for some $\sigma \geq 1$.
  $\Rightarrow$ Yields *well centered* point at expense of increasing infeasibility.
  $\Rightarrow$ Practical experience disappointing

- Adjusting for changed (upper) bounds
  $\Rightarrow$ No real progress, upper bounds are bad news for warmstarting

- Attempt to *unblock* blocking components (by perturbing μ-center).

$\Rightarrow$ Limited success for small problems, disappointing for large ones.

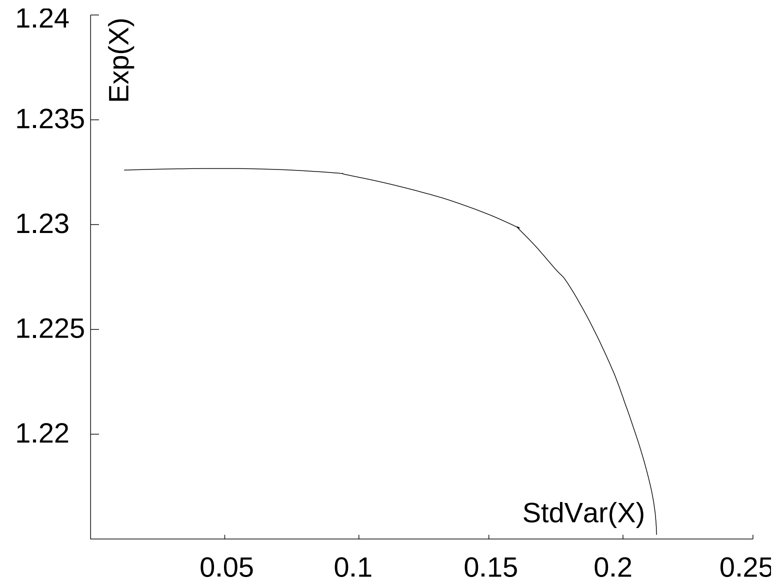
Application: Efficient Frontier

- Investors differ in their attitude to risk.
- Captured by the *risk-aversion* parameter ρ in the mean-variance model.
- Efficient Frontier provides a more complete understanding

solve

$$\max \mathbb{E}(X) - \rho \mathrm{Var}(X)$$

for a range of values for ρ



Gives max achievable expected gain for a given risk, or vice versa.

Efficient Frontier (Results)

Problem	constraints	variables	ρ =0.001	0.005	0.01	0.05	0.1	0.5	1	5	10
ALM11	223.321	76.881	14	14	14	14	14	13	17	16	17
			14	5	5	5	4	5	5	8	8
ALM12	533.725	198.525	14	14	14	14	14	15	18	18	17
			14	5	5	5	6	5	5	9	10
ALM13	5.982.604	16.316.191	24	23	24	23	25	22	24	23	24
			24	8	11	13	11	13	12	12	14
ALM14	70.575.308	192.478.111	52	53	45	43	44	42	44	46	46
			52	13	13	15	15	16	16	23	25

Applied strategies:

- Save ($\mu = 0.01$)-center for warmstarting
- re-center
- Adjust z for changes in objective

$\Rightarrow$ Warmstarting IPMs can be very effective

Towards a warmstarted SQP scheme

Generally considered a bad idea (for lack of warmstarts), but …

… if it wasn't, how would we do it?

Good ideas:

- Changes in Objective ✔
- Changes in Convexity ✔
  (additional $\gamma x^\top x$-term can be interpreted as regularization)

Bad ideas:

- Bounds/ Changes in Bounds ✘
  $\Rightarrow$ ($\|.\|_\infty$-) Trust Regions ✘
- Changes in problem structure (Phase I/Phase II) ✘

Proposed l_2-Interior Point-SQP Scheme

NLP Problem:

$$\begin{array}{rl} \min & f(x) \\ \text{s.t.} & g(x) = 0 \\ & x \geq 0 \end{array} \qquad \text{(NLP)}$$

The solution to the l_2-Trust Region QP subproblem:

$$\begin{array}{rl} \min & \nabla f(x)^\top d + \frac{1}{2} d^\top W d \\ \text{s.t.} & \nabla g(x)^\top d = -g(x) \\ & x + d \geq 0, \ \|d\|_2 \leq \rho \end{array}$$

is also a solution to

$$\begin{array}{rl} \min & \nabla f(x)^\top d + \frac{1}{2} d^\top [W + \delta I] d \\ \text{s.t.} & \nabla g(x)^\top d = -g(x) \\ & x + d \geq 0 \end{array}$$

for some value of $\delta = \delta(\rho)$

$\Rightarrow$ Idea: use δ rather than ρ as a trust region parameter

Proposed l_2-Interior Point-SQP Scheme II

Issues:

1. Trust Region: for $\rho \to 0 (\delta \to \infty)$ would expect $\|d\|_2 \to 0$.
   Problem if $d = 0$ is not feasible (no change when increasing δ).

2. For $k \to \infty$ would expect TR constraint to not be active.
   Problem: "Trust Region" is always active at solution of QP.

Proposed solutions:

1. use *feasibility formulation*

2. expect $\delta \to 0$ for $k \to \infty$
   $\Rightarrow$ SQP should not be disturbed to much.
   A small *regularizing* term δI is beneficial for IPM.

Solving QP by Interior Point Method

$$\min c^\top x + \tfrac{1}{2} x^\top Q x \quad \text{s.t.} \quad \begin{array}{rcl} Ax & = & b \\ x & \geq & 0 \end{array}$$

Optimality conditions:

$$\begin{array}{rcl} c + Qx - A^\top y - z & = & 0 \\ Ax & = & b \\ XZe & = & 0 \quad (\mu e) \\ x, z & \geq & 0 \end{array}$$

$\Rightarrow$ Newton Step:

$$\begin{bmatrix} -Q - \Theta & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} Qx + c - A^\top y - \mu X^{-1} e \\ b - Ax \end{bmatrix}$$

where

$$\Theta = X^{-1} Z, \quad X = \mathrm{diag}(x), \quad Z = \mathrm{diag}(z)$$

Solving feasibility QP by Interior Point Method

$$\min c^\top x + \tfrac{1}{2}x^\top Qx + \tfrac{1}{2}\gamma s^\top s \;\; \text{s.t.} \quad \begin{aligned} Ax + s &= b \\ x &\geq 0 \end{aligned}$$

Optimality conditions:

$$\begin{aligned} c + Qx - A^\top y - z &= 0 \\ Ax + \tfrac{1}{\gamma} y &= b \\ XZe &= 0 \quad (\mu e) \\ x, z &\geq 0 \end{aligned}$$

$\Rightarrow$ Newton Step:

$$\begin{bmatrix} -Q - \Theta & A^\top \\ A & \frac{1}{\gamma} I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} Qx + c - A^\top y - \mu X^{-1} e \\ b - Ax - \frac{1}{\gamma} y \end{bmatrix}$$

where

$$\Theta = X^{-1}Z, \quad X = \mathrm{diag}(x), \quad Z = \mathrm{diag}(z)$$

Proposed l_2-Interior Point-SQP Scheme III

NLP Problem:

$$\begin{array}{rl} \min & f(x) \\ \text{s.t.} & g(x) = 0 \\ & x \geq 0 \end{array} \qquad \text{(NLP)}$$

Phase-2 QP:

$$\begin{array}{rl} \min & \nabla f(x)^\top d + \frac{1}{2} d^\top W d \\ \text{s.t.} & \nabla g(x)^\top d = -g(x) \\ & x + d \geq 0 \end{array} \qquad \text{(QP-2)}$$

Phase-1 QP (infeasibility):

$$\begin{array}{rl} \min & \nabla f(x)^\top d + \frac{1}{2} d^\top W d + \gamma s^\top s \\ \text{s.t.} & \nabla g(x)^\top d + s = -g(x) \\ & x + d \geq 0 \end{array} \qquad \text{(QP-1)}$$

Both (QP-1) and (QP-2) use the same structure for the IPM Newton-Step.
$\Rightarrow$ Can be warmstarted from each other.

Proposed l_2-Interior Point-SQP Scheme IV

```
set phase=2   ; choose x, λ
loop
    solve (QP-phase) ⇒ get d, new λ
    if (Step acceptable)
        x ← x + d
        maybe increase TR (decrease δ)
    else
        decrease TR (increase δ)
        if (previous decrease had no effect)
            phase  = 1, γ = 5
    end if
    in phase-1:
        (possibly) increase γ
        if (constraints almost satisfied)
            phase  = 2
        end if
end loop
```

Future Work:

- HPCx
- Robustify SQP scheme and investigate success of warmstarting strategies
- Solve QPs to approximate optimality
- Link to modelling language (SMPS is supported)
- Implement other structures/other strategies (Iterative solver etc)