Numeric and Optimize Method

1. Matrix Decomposition/ Factorization

reference:

- wiki

1.1. LU

reference:

- wiki

fomulation:

$$A = LU \tag{<1>}$$

- L: lower triangular
- U: upper triangular

1.1.1. Partial Piviting

+fomulation:

$$PA = LU \tag{<2>}$$

- P: permutation that reorder rows

feature:

- numerically stable

1.1.2. Full Piviting

+fomulation:

$$PAQ = LU \tag{<3>}$$

- Q: permutation that reorder columns

1.1.3. LDU

+fomulation:

$$A = LDU \qquad <4>$$

- D: diagonal
- L,U: +unitraingular

1.2. Cholesky

formulation:

$$A = UU^T \qquad <5>$$

- A: symmetric, positive (semi-)defined
- U: upper traingular, (semi-)positive diagonal entries

1.2.1. LDL/LDLT

+formulation:

$$A = LDL^T \qquad <6>$$

- L: lower unitraingular
- D: diagonal

feature:

- +square-root-free

1.3. QR

formulation:

$$A = QR \qquad <7>$$

- Q: orthogonal
- R: upper traingular

compute:

1.3.1. Gram-Schmidt Process

feature:
- low numeric stability
- easy implementation

1.3.2. Householder Reflections

feature
- better numeric stability than Gram-Schmidt Process
- bandwidth heavy
- not parallelizable

1.3.3. Givens Rotations

feature:
- sparse
- parallelizable

2. Equaltion

2.1. Linear

formulation:

$$A\boldsymbol{x} + \boldsymbol{b} = 0 \tag*{<8>}$$

transform:
if $\operatorname{rank}(A) < \boldsymbol{b}$

$$A^\top A\boldsymbol{x} + A^\top \boldsymbol{b} = 0 \tag{<9>}$$

$$\begin{gathered} A'\boldsymbol{x} + \boldsymbol{b}' = 0 \\ \to <8> \end{gathered} \tag{<10>}$$

2.1.1. Conjugate Gradient

reference:

- wiki
- cornell.edu

2.2. Continuous

formulation:

$$\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{b} = 0 \tag{<11>}$$

2.2.1. Newton

transform:

$$\begin{gathered} \nabla_{\boldsymbol{x}_k} \boldsymbol{f}(\boldsymbol{x}_k)(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) + \boldsymbol{f}(\boldsymbol{x}_k) + b = 0 \\ \to <8> \end{gathered} \tag{<12>}$$

3. Optimization

formulation:

$$\min_x f(x) \tag{<13>}$$

3.1. Quadral

formulation:

$$\min_{\boldsymbol{x}} \boldsymbol{x}^\top A\boldsymbol{x} + \boldsymbol{b}\boldsymbol{x} \tag*{<14>}$$

or

$$\min_{\boldsymbol{x}} \| A\boldsymbol{x} + \boldsymbol{b} \|_2 \tag*{<15>}$$

transform:

$$\begin{gathered} A\boldsymbol{x} + \boldsymbol{b} = 0 \\ \to <8> \end{gathered} \tag*{<16>}$$

or

$$2A^\top(A\boldsymbol{x} + \boldsymbol{b}) = 0 \tag*{<17>}$$

$$A^\top A\boldsymbol{x} + A^\top\boldsymbol{b} = 0 \tag*{<18>}$$

$$\begin{gathered} A'\boldsymbol{x} + \boldsymbol{b}' = 0 \\ \to <8> \end{gathered} \tag*{<19>}$$

3.2. Continuous

formulation:

$$\begin{gathered} \min_{\boldsymbol{x}} f(\boldsymbol{x}), \\ \nabla_{\boldsymbol{x}} f(\boldsymbol{x}_k) = \boldsymbol{g}, \nabla^2_{\boldsymbol{x}} f(\boldsymbol{x}_k) = H \end{gathered} \tag*{<20>}$$

3.2.1. Newton

transform:

$$
\begin{gathered}
f(\boldsymbol{x})_k = \boldsymbol{x}^\top H \boldsymbol{x}_k + \boldsymbol{g}\boldsymbol{x}_k, \\
\min_{\boldsymbol{x}_{k+1}} f(\boldsymbol{x})_k \qquad <21> \\
\to <15>
\end{gathered}
$$

3.2.2. Quasi-Newton

reference:

- wiki

3.2.2.1. BFGS

reference:

- wiki

3.2.2.1.1. L-BFGS

reference:

- wiki

3.2.2.2. Compact Representation

reference:

- wiki

4. Constraint

4.1. Equality

formulation:

$$
\boldsymbol{n}\boldsymbol{x} = 0 \qquad <22>
$$

kkt:

$$
\begin{gathered}
\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) + \nabla_{\boldsymbol{x}} \boldsymbol{n}(\boldsymbol{x}) \boldsymbol{\lambda} = 0, \\
\boldsymbol{n}(\boldsymbol{x}) = 0
\end{gathered}
\tag{<23>}
$$

4.1.1. Linear Equality

formulation:

$$
\begin{gathered}
N \boldsymbol{x} + \boldsymbol{m} = 0, \\
\operatorname{rank}(N^\top) < \dim(\boldsymbol{x})
\end{gathered}
\tag{<24>}
$$

or

$$
\begin{gathered}
\boldsymbol{x} = N \boldsymbol{\lambda}, \\
\operatorname{rank}(N) < \dim(\boldsymbol{x})
\end{gathered}
\tag{<25>}
$$

4.1.1.1. Qualdral Optimization (Linear Least Squares)

reference:

- wiki

transform:

$$
\begin{gathered}
<14>, <24>, <23> \Rightarrow \\
N \boldsymbol{\lambda} + 2A \boldsymbol{x} + \boldsymbol{b} = 0
\end{gathered}
\tag{<26>}
$$

$$
\begin{gathered}
<24>, <26> \Leftrightarrow \\
\begin{pmatrix} 2A & N \\ N & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{\lambda} \end{pmatrix} + \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{m} \end{pmatrix} = 0
\end{gathered}
\tag{<27>}
$$

$$
\begin{gathered}
A' \boldsymbol{x}' + \boldsymbol{b}' = 0 \\
\rightarrow <8>
\end{gathered}
\tag{<28>}
$$

or

$$
\begin{gathered}
<15>, <25> \Rightarrow \\
\min_{\boldsymbol{\lambda}} \| \, A N \boldsymbol{\lambda} - \boldsymbol{b} \, \|_2
\end{gathered} \tag{<29>}
$$

$$
\begin{gathered}
\min_{\boldsymbol{x}'} \| \, A' \boldsymbol{x}' - \boldsymbol{b} \, \|_2 \\
\to <14>
\end{gathered} \tag{<30>}
$$

4.2. InEquality

formulation:

$$
\boldsymbol{n}(\boldsymbol{x}) \geq 0 \tag{<31>}
$$

kkt:

$$
\begin{gathered}
\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) + \nabla_{\boldsymbol{x}} \boldsymbol{n}(\boldsymbol{x}) \boldsymbol{\lambda} = 0, \\
\boldsymbol{\lambda} \boldsymbol{n}(\boldsymbol{x}) = 0, \\
\boldsymbol{\lambda} \geq 0
\end{gathered} \tag{<32>}
$$

4.2.1. Interior Point (Barrier)

reference:

- wiki

4.2.1.1. Path Following

transform:

$$
\begin{gathered}
\min_{\boldsymbol{x}_{t_k} = \boldsymbol{x}} f(\boldsymbol{x}) + t_k \boldsymbol{n}'(\boldsymbol{x}), t_{k+1} = \theta t_k \\
\to <13>
\end{gathered} \tag{<33>}
$$

4.2.1.2. Primal Dual

reference:

- cmu.edu

compare:

- little slower than Path Following but more accurate.

transform:

$$\begin{gathered} <33>, \\ n'(\boldsymbol{x}) := \| -\mathbf{log}(\boldsymbol{n}(\boldsymbol{x})) \|_1 \end{gathered} \tag{<34>}$$

$$\begin{gathered} \min_{\boldsymbol{x}_{t_k}=\boldsymbol{x}} f(\boldsymbol{x}) + t_k \| -\mathbf{log}(\boldsymbol{n}(\boldsymbol{x})) \|_1, \\ \boldsymbol{\lambda} \odot \boldsymbol{n}(x) = t_k \mathbf{1} \end{gathered} \tag{<35>}$$

$$\begin{gathered} <35> \Rightarrow \\ \nabla_{\boldsymbol{x}} f(\boldsymbol{x}) + t_k \nabla_{\boldsymbol{x}} \boldsymbol{n}(\boldsymbol{x}) \left(-\boldsymbol{n}(\boldsymbol{x})^{\bigcirc -\mathbf{1}}\right) = 0 \end{gathered} \tag{<36>}$$

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) + t_k \nabla_{\boldsymbol{x}} \boldsymbol{n}(\boldsymbol{x}) \left(-t_k^{-1} \boldsymbol{\lambda}\right) = 0 \tag{<37>}$$

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) - \nabla_{\boldsymbol{x}} \boldsymbol{n}(\boldsymbol{x}) \boldsymbol{\lambda} = 0 \tag{<38>}$$

$$\begin{gathered} <38>, <35> \Rightarrow \\ \nabla_{\boldsymbol{x}} f(\boldsymbol{x}) - \nabla_{\boldsymbol{x}} \boldsymbol{n}(\boldsymbol{x}) \boldsymbol{\lambda} = 0, \\ \boldsymbol{\lambda} \odot \boldsymbol{n}(x) = t_k \mathbf{1} \\ \to <11> \end{gathered} \tag{<39>}$$

or

$$
\begin{gathered}
\lt 32\gt \rightarrow \\
\boldsymbol{x}_k := \boldsymbol{x}, \\
\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) + \nabla_{\boldsymbol{x}} \boldsymbol{n}(\boldsymbol{x}) \boldsymbol{\lambda} = 0, \\
\boldsymbol{\lambda} \boldsymbol{n}(\boldsymbol{x}) = t_k, \\
t_{k+1} = \theta t_k \\
\rightarrow \lt 11\gt
\end{gathered}
\tag*{$\lt 40\gt$}
$$

4.2.2. Linear InEquality

formulation:

$$
N\boldsymbol{x} + \boldsymbol{m} \geq 0 \tag*{$\lt 41\gt$}
$$