Quadratic Programming

January 1, 2023

Quadratic Programming

General Formulation

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x}, \text{ s.t. } \mathbf{b}_l \leq \mathbf{A}\mathbf{x} \leq \mathbf{b}_u \text{ and } \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \tag{1}$$

- $\mathbf{b}_l \leq \mathbf{A}\mathbf{x} \leq \mathbf{b}_u$ are general linear constraints
- $\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}$ are simple bound constraints
- problem (1) is convex QP if Hessian matrix $\mathbf{H}$ is positive definite, otherwise it is a general QP problem.
- Mathwrist takes the general form (1).
- without loss of generality, we will be looking at a convenient form,

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x}, \text{ s.t. } \mathbf{A}\mathbf{x} \geq \mathbf{b} \tag{2}$$

Quadratic Programming

Review of Active Set Method

- maintain index sets $\mathcal{W}$ and $\mathcal{N}$ for working and non-working general constraints respectively.
- seek a descent null space direction $\mathbf{p} = \mathbf{Z}\mathbf{p}_z$ for some $\mathbf{p}_z$, where $\mathbf{Z}$ is the null space columns of the $\mathbf{QR}$ factorization

$$\mathbf{A}_{\mathcal{W}}^T = \underbrace{\left(\mathbf{Y} \mid \mathbf{Z} \right)}_{\mathbf{Q}} \begin{pmatrix} \mathbf{R} \\ 0 \end{pmatrix}$$

- at a non-stationary iteration k, make a step move $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \mathbf{p}$, where α is determined by the first blocking constraint a_i. i is then added to working set $\mathcal{W}$.

Quadratic Programming

Convex QP: KKT System

We solve the following KKT system for search direction **p** and Lagrange multipliers λ,

$$\begin{pmatrix} \mathbf{H} & \mathbf{A}_{\mathcal{W}}^T \\ \mathbf{A}_{\mathcal{W}} & \end{pmatrix} \begin{pmatrix} -\mathbf{p} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{g}(\mathbf{x}_k) \\ \mathbf{0} \end{pmatrix} \tag{3}$$

The second KKT equation implies **p** is a null space direction wrt the set of working constraints.

Quadratic Programming

Convex QP: Null Space Direction

- write $\mathbf{p} = \mathbf{Z}\mathbf{p}_z$ for some $\mathbf{p}_z$.
- the first KKT equation in (3) is

$$-\mathbf{H}\mathbf{Z}\mathbf{p}_z + \mathbf{A}_{\mathcal{W}}^T\lambda = \mathbf{g}(\mathbf{x}_k) \tag{4}$$

- multiply $\mathbf{Z}^T$ at both sides of (4), solve $\tilde{\mathbf{H}}\mathbf{p}_z = -\tilde{\mathbf{g}}$ where
  - $\tilde{\mathbf{H}} = \mathbf{Z}^T\mathbf{H}\mathbf{Z}$ is the reduce Hessian.
  - $\tilde{\mathbf{g}} = \mathbf{Z}^T\mathbf{g}(\mathbf{x}_k)$ is the reduced gradient.

Quadratic Programming

Convex QP: Performance Efficiency

- retain **QR** factorization of $\mathbf{A}_{\mathcal{W}}^T$.
- retain Cholesky factorization of reduced Hessian $\tilde{\mathbf{H}} = \mathbf{L}\mathbf{L}^T$ to solve $\mathbf{p}_z$ in equation (4).
- $\mathbf{A}_{\mathcal{W}}$ and $\tilde{\mathbf{H}}$ changes whenever the working set $\mathcal{W}$ changes.
- apply economical updates on **QR** and Cholesky instead of refactoring the new matrix from the scratch.

Quadratic Programming

Convex QP: Step Update

- move along $\mathbf{p}$ at step length $\alpha \in [0, 1]$
- if a blocking constraint a_i is hit before making a unit step move $\alpha = 1$,
  - add constraint i to $\mathcal{W}$.
  - update $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \mathbf{p}$ and start the $k + 1$ iteration.
- otherwise, making a unit step move at $\alpha = 1$ reaches the local optimal, test Lagrange multiplier conditions

Quadratic Programming

Convex QP: Lagrange Multipliers

- after $\mathbf{p}$ is computed, from the first KKT equation in (5) we have

$$\mathbf{A}_{\mathcal{W}}^T \lambda = \mathbf{g}(\mathbf{x}_k) + \mathbf{H}\mathbf{p}$$

- multiply $\mathbf{Y}^T$ at both sides and solve

$$\mathbf{R}\lambda = \mathbf{Y}^T(\mathbf{g}(\mathbf{x}_k) + \mathbf{H}\mathbf{p})$$

Quadratic Programming

Convex QP: Local Optimal

- when a local optimal is reached, the algorithm terminates if $\lambda_i \geq 0$ for all lower bounded constraints $i \in \mathcal{W}$.
- if lower bounded constraint $i \in \mathcal{W}$ has the most negative Lagrange multiplier, delete i from $\mathcal{W}$ and add it to $\mathcal{N}$.
- such active constraint deletion operation results to a descent feasible direction in the next iteration.

Quadratic Programming

Handle Simple Bound Constraints

- partion variable $\mathbf{x}$ to a set of fixed variables $\mathbf{x}_b$ and the complement set of free variables $\mathbf{x}_f$.
- $\mathbf{x}_b$ have values fixed at one side of the bounds, effectively as active constraints $\mathbf{I}\mathbf{x}_b = \mathbf{v}_b$ for relevant boundary value $\mathbf{v}_b$.
- KKT system (3) then has the following block structure,

$$\begin{pmatrix} \mathbf{H}_{bb} & \mathbf{H}_{bf} & \mathbf{I} & \mathbf{B}_{\mathcal{W}}^T \\ \mathbf{H}_{fb} & \mathbf{H}_{ff} & & \mathbf{F}_{\mathcal{W}}^T \\ \mathbf{I} & & & \\ \mathbf{B}_{\mathcal{W}} & \mathbf{F}_{\mathcal{W}} & & \end{pmatrix} \begin{pmatrix} -\mathbf{p}_b \\ -\mathbf{p}_f \\ \lambda_b \\ \lambda_{\mathcal{W}} \end{pmatrix} = \begin{pmatrix} \mathbf{g}_b \\ \mathbf{g}_f \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{5}$$

- subscript b and f represent the sub matrix/vector blocks that correspond to fixed and free variables respectively.
- the 3rd equation $\mathbf{I}\mathbf{p}_b = \mathbf{0} \rightarrow \mathbf{p}_b = 0$, representing $\mathbf{x}_b$ is fixed.
- seek a null space direction $\mathbf{p}_f$ wrt free variables $\mathbf{x}_f$.

Quadratic Programming

General QP: What is special at an active set iteration?

- need efficiently determine the definiteness of $\tilde{\mathbf{H}}$, along the course of active set $\mathcal{W}$ changes.
- cases to consider at iteration k,
  - $\tilde{\mathbf{H}}$ is indefinite, surely not optimal.
  - $\tilde{\mathbf{H}}$ is positive semi-definite and $\mathbf{x}_k$ is not a stationary point wrt $\mathcal{W}$, not optimal either.
  - $\tilde{\mathbf{H}}$ is at least positive semi-definite, $\mathbf{x}_k$ is a stationary point wrt $\mathcal{W}$, it is a local optimal.

Quadratic Programming

General QP: Local Optimal

At a local optimal point, everything stays same as the convex QP situation. The active set algorithm will

- compute Lagrange multipliers from KKT system (3).
- test Lagrange multiplier conditions, $\lambda_{\mathcal{W}} \geq 0$ for lower bounded constraints, ≤ 0 for upper bounded ones.
- terminate if $\lambda_{\mathcal{W}}$ pass the test.
- otherwise, find the worst $i \in \mathcal{W}$ and delete i from $\mathcal{W}$.

Quadratic Programming

General QP: Inertia Control and Extended KKT

- ensure $\tilde{\mathbf{H}}$ has at most one negative eigen value.
- at a local optimal, if deleting an active constraint a_* will cause the new $\tilde{\mathbf{H}}$ being indefinite, mark a_* as "pending deletion" and still keep it in an extended KKT system.

$$\begin{pmatrix} \mathbf{H} & \mathbf{A}_{\mathcal{W}}^T & a_*^T \\ \mathbf{A}_{\mathcal{W}} & & \\ a_* & & \end{pmatrix}$$

- keep moving along descent feasible directions and adding blocking constraints until $\tilde{\mathbf{H}}$ becomes positive definite again, at that point, physically remove a_* from the extended KKT system.

Quadratic Programming

General QP: p for Indefinite $\tilde{\mathrm{H}}_{ff}$

- if $\mathbf{x}_k$ is non-stationary wrt $\mathcal{W}$, **p** is a direction of negative curvature, solved from KKT

$$\begin{pmatrix} \mathbf{H} & \mathbf{A}_{\mathcal{W}}^T \\ \mathbf{A}_{\mathcal{W}} & \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \nu \end{pmatrix} = \begin{pmatrix} \mathbf{g}(\mathbf{x}_k) \\ \mathbf{0} \end{pmatrix} \qquad (6)$$

- if stationary, a negative curvature direction **p** can be found from extended KKT

$$\begin{pmatrix} \mathbf{H} & \mathbf{A}_{\mathcal{W}}^T & a_*^T \\ \mathbf{A}_{\mathcal{W}} & & \\ a_* & & \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \nu \\ \nu_* \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ 1 \end{pmatrix} \qquad (7)$$

Quadratic Programming

General QP: p for Semi-definite and Singular $\tilde{\mathbf{H}}$

- $\mathbf{p}$ is a direction of zero curvature solved from KKT

$$\begin{pmatrix} \mathbf{H} & \mathbf{A}_{\mathcal{W}}^T \\ \mathbf{A}_{\mathcal{W}} & \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \nu \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{8}$$

Quadratic Programming

Initial Working Set and Feasible Point

- find an initial feasible point same as linear programming active set method.
- inertial control requires to start with a positive definite $\tilde{\mathbf{H}}$.
- for general QP, perform a partial Cholesky [5],

$$\mathbf{P}^T\tilde{\mathbf{H}}\mathbf{P} = \begin{pmatrix} \mathbf{L}_{11} & \\ \mathbf{L}_{21} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{D} & \\ & \mathbf{K} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{11}^T & \mathbf{L}_{21}^T \\ & \mathbf{I} \end{pmatrix}$$

  , where permutation $\mathbf{P}$ arranges the null space as $\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_+ & \mathbf{Z}_- \end{pmatrix}$ such that $\mathbf{Z}_+^T\mathbf{H}\mathbf{Z}_+ = \mathbf{L}_{11}\mathbf{D}\mathbf{L}_{11}^T$ is positive definite.
- $\mathbf{Z}_-$ are treated as artificial constraints and added to the initial working set.

Quadratic Programming

General QP: Performance Efficiency

- same matrix factorization update technique as convex QP.
- stationary condition and definiteness of $\tilde{\mathbf{H}}$ is determined by various vector update schemes, details in [3].
- KKT solutions are economically updated as well, details in [3]

References I

[1] Jorge Nocedal and Stephen J. Wright: Numerical Optimization, Springer, 1999

[2] Philip E. Gill, Walter Murray and Margaret H. Wright: Practical Optimization, Academic Press, 1981

[3] Philip E. Gill, Walter Murray, Michael A. Saunders and Margaret H. Wright: Inertia-Controlling Methods For General Quadratic Programming, SIAM Review, Volume 33, Number 1, 1991, pages 1-36.

[4] Philip E. Gill, Walter Murray, Michael A. Saunders and Margaret H. Wright: Procedures for Optimization Problems with a Mixture of Bounds and General Linear Constraints, ACM Transactions on Mathematical Software, Vol.10, No. 3, September 1984, Pages 282-298

References II

[5] Anders. Forsgren, Philip E. Gill and Walter Murray: Computing Modified Newton Directions Using a Partial Cholesky Factorization, SIAM J. SCI. COMPUT. Vol. 16, No. 1, pp. 139-150