Quadratic Optimization

Introduction and Equality Constrained QP

John Bagterp Jørgensen

Department of Applied Mathematics and Computer Science
The Technical University of Denmark
jbjo@dtu.dk

02612 Constrained Optimization
March 2020

Outline

Convex Quadratic Program

The convex quadratic programming problem

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & F(x) = \frac{1}{2} x' H x + g' x + \gamma && \text{(1a)} \\
s.t. \quad & a_i' x = b_i \qquad i = 1, \ldots, m_a && \text{(1b)} \\
& c_i' x \geq d_i \qquad i = 1, \ldots, m_c && \text{(1c)}
\end{aligned}$$

H positive semi-definite: Convex QP
H positive definite: Strictly Convex QP
Convex QP

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & F(x) = \frac{1}{2} x' H x + g' x + \gamma && \text{(2a)} \\
s.t. \quad & A' x = b && \text{(2b)} \\
& C' x \geq d && \text{(2c)}
\end{aligned}$$

Equality Constrained QP

Equality constrained QP

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} x' H x + g' x + \gamma \tag{3a}$$

$$s.t. \quad a_i' x = b_i \qquad i = 1, 2, \ldots, m_a \tag{3b}$$

Equality constrained QP in matrix form

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} x' H x + g' x + \gamma \tag{4a}$$

$$s.t. \quad A' x = b \tag{4b}$$

$A = \begin{bmatrix} a_1 & a_2 & \ldots & a_{m_a} \end{bmatrix}$ and $b = \begin{bmatrix} b_1 & b_2 & \ldots & b_{m_a} \end{bmatrix}'$
The associated necessary and sufficient conditions for optimality

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix} \tag{5}$$

Applications of Quadratic Optimization/Programming

Applications of convex QPs:

1. Return-risk trade-off in portfolio-optimization (Markowitz portfolio optimization problem)
2. Constrained least squares regression
3. Subproblem in sequential quadratic programming (SQP) for solution of nonlinear programming problems
4. Discrete-time optimal control: Model predictive control (MPC), moving-horizon estimation (MHE)
5. Dynamic versions of Markowitz portfolio optimization (MPC)
6. Discrete-time optimal control: Production-Inventory models
7. Huber regression of linear models
8. De-trending (ℓ_1-detrending)

Example: Portfolio Optimization

- $i = 1, 2, \ldots, n$ assets to invest in
- x_i fraction of fund invested in asset i
- r_i return of asset i. $r \in \mathbb{F}(\mu, H)$

Return, expected return and variance of the portfolio

$$R = \sum_{i=1}^{n} r_i x_i = r'x \tag{6}$$

$$\bar{R} = E\{R\} = E\{r'x\} = \mu' x \tag{7}$$

$$V\{R\} = E\left\{(R - \bar{R})^2\right\} = E\{x'(r-\mu)(r-\mu)'x\} = x'Hx \tag{8}$$

Markowitz Portfolio Optimization Problem (risk-tolerance parameter $\kappa \in [0, \infty)$):

$$\max_{x \in \mathbb{R}^n} \quad F(x) = x'\mu - \kappa x' H x \tag{9a}$$

$$s.t. \quad \sum_{i=1}^{n} x_i = 1 \tag{9b}$$

$$x \geq 0 \tag{9c}$$

Example: Portfolio Optimization

Markowitz Portfolio Optimization Problem ($\kappa \in [0, \infty)$):

$$\max_{x \in \mathbb{R}^n} \quad F(x) = x'\mu - \kappa x' H x \tag{10a}$$

$$s.t. \quad \sum_{i=1}^{n} x_i = 1 \tag{10b}$$

$$x \geq 0 \tag{10c}$$

Alternative formulation (that does not always has a feasible solution)

$$\max_{x \in \mathbb{R}^n} \quad F(x) = x'\mu \tag{11a}$$

$$s.t. \quad x' H x \leq \gamma \tag{11b}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{11c}$$

$$x \geq 0 \tag{11d}$$

Example: Portfolio Optimization

Markowitz Portfolio Optimization Problem ($\kappa \in [0, \infty)$):

$$\max_{x \in \mathbb{R}^n} \quad F(x) = x'\mu - \kappa x' H x \tag{12a}$$

$$s.t. \quad \sum_{i=1}^{n} x_i = 1 \tag{12b}$$

$$x \geq 0 \tag{12c}$$

Alternative formulation $R \in \left[\min_i r_i \quad \max_i r_i\right]$

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} x' H x \tag{13a}$$

$$s.t. \quad \mu' x = R \tag{13b}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{13c}$$

$$x \geq 0 \tag{13d}$$

Example: Constrained Least Squares Regression

Constrained Least Squares Regression Problem

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} \|Ax - b\|_2^2 \tag{14a}$$

$$s.t. \qquad l \leq x \leq u \tag{14b}$$

The objective function is quadratic

$$\begin{aligned} F(x) &= \frac{1}{2} \|Ax - b\|_2^2 = \frac{1}{2}(Ax - b)'(Ax - b) \\ &= \frac{1}{2} x' A' A x + (-A'b)'x + \frac{1}{2} b'b = \frac{1}{2} x' H x + g' x + \gamma \end{aligned} \tag{15}$$

with

$$H = A'A \quad g = -A'b \quad \gamma = \frac{1}{2} b'b \tag{16}$$

Consequently, it is a convex quadratic program

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} x' H x + g' x + \gamma \tag{17a}$$

$$s.t. \qquad l \leq x \leq u \tag{17b}$$

Example: Constrained Weighted Least Squares Regression

Constrained Least Squares Regression Problem

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} \|Ax - b\|_W^2 \tag{18a}$$

$$s.t. \qquad l \leq x \leq u \tag{18b}$$

The objective function is quadratic

$$\begin{aligned} F(x) &= \frac{1}{2} \|Ax - b\|_W^2 = \frac{1}{2}(Ax - b)' W (Ax - b) \\ &= \frac{1}{2} x' A' W A x + (-A' W b)' x + \frac{1}{2} b' W b = \frac{1}{2} x' H x + g' x + \gamma \end{aligned} \tag{19}$$

with

$$H = A' W A \quad g = -A' W b \quad \gamma = \frac{1}{2} b' W b \tag{20}$$

Consequently, it is a convex quadratic program

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} x' H x + g' x + \gamma \tag{21a}$$

$$s.t. \qquad l \leq x \leq u \tag{21b}$$

Example: Linear Model Predictive Control

Compute $\{u_k\}_{k=0}^{N-1}$ such that the predicted output trajectory $\{y_k\}_{k=0}^{N-1}$ follows the specified output trajectory $\{r_k\}_{k=0}^{N-1}$ as good as possible.

$$\min \quad \phi = \frac{1}{2} \sum_{k=0}^{N-1} \|y_k - r_k\|_Q^2 + \|\Delta u_k\|_S^2 \tag{22a}$$

$$s.t. \quad x_{k+1} = Ax_k + Bu_k + Ed_k \qquad k = 0, 1, \ldots, N-1 \tag{22b}$$

$$y_k = Cx_k + c_k \qquad k = 0, 1, \ldots, N-1 \tag{22c}$$

$$u_{\min} \leq u_k \leq u_{\max} \qquad k = 0, 1, \ldots, N-1 \tag{22d}$$

$$\Delta_l \leq \Delta u_k \leq \Delta_u \qquad k = 0, 1, \ldots, N-1 \tag{22e}$$

$$\Delta u_k = u_k - u_{k-1} \tag{23}$$

Constrained least squares problem with dynamic structure.

Example: Linear Moving Horizon Estimation

- Measurements $\{y_k\}_{k=0}^N$ given
- Exogeneous signals $\{b_k\}_{k=0}^N$ and $\{c_k\}_{k=0}^N$ given
- Determine the states $\{x_k\}_{k=0}^N$, the process noise $\{w_k\}_{k=0}^N$, and the measurement noise $\{v_k\}_{k=0}^N$

To do this solve

$$\min_{\{x_k, w_k, v_k\}} \quad \phi = \frac{1}{2} \|x_0 - \bar{x}_0\|_{P_0^{-1}}^2 + \frac{1}{2} \sum_{k=0}^{N} \|w_k\|_{Q^{-1}}^2 + \|v_k\|_{R^{-1}}^2 \tag{24a}$$

$$s.t. \quad x_{k+1} = Ax_k + Hw_k + b_k \qquad k = 0, 1, \ldots, N \tag{24b}$$

$$y_k = Cx_k + c_k + v_k \qquad k = 0, 1, \ldots, N \tag{24c}$$

$$w_{\min} \leq w_k \leq w_{\max} \qquad k = 0, 1, \ldots, N \tag{24d}$$

$$x_{\min} \leq x_k \leq x_{\max} \qquad k = 0, 1, \ldots, N \tag{24e}$$

This is a smoothing problem, and an example of a constrained least squares regression problem.

Example: Production-Inventory Optimization

- u_k change in labor force (hiring / firing) in period k
- L_k total labor force at time k
- d_k demand in period k. Known (estimated)
- I_k inventory at time k
- p number of units produced per worker per time period
- b production capacity of machinery (units / period)

$$
\begin{aligned}
\min_{\{u_k\}_{k=0}^{N-1}} \quad & \phi = \sum_{k=0}^{N-1} \left(c_1 u_k^2 + c_2 I_k \right) && && \text{(25a)} \\
s.t. \quad & L_{k+1} = L_k + u_k && k = 0, 1, \ldots, N-1 && \text{(25b)} \\
& I_{k+1} = I_k + pL_k - d_k && k = 0, 1, \ldots, N-1 && \text{(25c)} \\
& 0 \leq L_k \leq b/p && && \text{(25d)} \\
& 0 \leq I_k && && \text{(25e)}
\end{aligned}
$$

Example: Huber Regression of Linear Models

Gertz and Wright (2001)

$$\min_{x \in \mathbb{R}^n} \quad \phi = \sum_{i=1}^{m} \rho\left([Ax - b]_i\right) \qquad \rho(t) = \begin{cases} \frac{1}{2}t^2 & |t| \leq \tau \\ \tau|t| - \frac{1}{2}\tau^2 & |t| > \tau \end{cases} \tag{26}$$

This problem can be expressed as (Mangasarian and Musicant, 2000)

$$\min_{x,y,z,w} \quad \phi = \frac{1}{2}w'w + \tau e'(y + z) \tag{27a}$$

$$s.t. \quad w - Ax + b - y + z = 0 \tag{27b}$$

$$y \geq 0 \tag{27c}$$

$$z \geq 0 \tag{27d}$$

or it can be obtained through solution of (Li and Swetits, 1998)

$$\min_{w} \quad \phi = \frac{1}{2}w'w + b'w \tag{28a}$$

$$s.t. \quad A'w = 0 \tag{28b}$$

$$-\tau e \leq w \leq \tau e \tag{28c}$$

Quadratic Unconstrained Optimization

Quadratic unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} x' H x + g' x + \gamma \tag{29}$$

Optimality condition (necessary and sufficient)

$$\nabla F(x) = Hx + g = 0 \qquad \Leftrightarrow \qquad Hx = -g \tag{30}$$

The optimum is

$$x = -H^{-1} g \tag{31}$$

Equality Constrained Convex QP

Equality constrained convex QP

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2}x'Hx + g'x + \gamma \tag{32a}$$

$$s.t. \qquad A'x = b \tag{32b}$$

Lagrangian

$$L(x, y) = \frac{1}{2}x'Hx + g'x + \gamma - y'(A'x - b) \tag{33}$$

First order necessary and sufficient optimality conditions

$$\nabla_x L(x, y) = Hx + g - Ay = 0 \tag{34a}$$

$$\nabla_y L(x, y) = -(A'x - b) = 0 \tag{34b}$$

This is the augmented equation (KKT equation)

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix} \tag{35}$$

Equality Constrained Convex QP

Equality constrained convex QP

$$\min_{x \in \mathbb{R}^n} \quad F(x) = \frac{1}{2} x' H x + g' x + \gamma \tag{36a}$$

$$s.t. \quad A' x = b \tag{36b}$$

The corresponding KKT-system

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix} \tag{37}$$

(KKT system, equilibrium system, augmented system)

Take home message:

Solution of the equality constrained convex QP corresponds to solution of a KKT-system.
Solution of a KKT-system corresponds to solution of an equality constrained convex QP.

Direct Solution of the KKT System

KKT system

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix} \tag{38}$$

- Matlab: General sparse matrix, dense matrix LU-factorization. Not tailored for this system.
- LAPACK: DGESV (general dense LU-solver), DSYSV (sense solver for symmetric indefinite systems, Bunch-Kaufman pivoting)
- Harwell Subroutine Library (HSL): MA27, MA57 for sparse symmetric indefinite matrices. MA48 for sparse general matrices.
- Problems with special structure: Dynamic problems (linear optimal control problems), support vector machines, ...

Explicit Expression for the Inverse of the KKT-matrix

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix}^{-1} = \begin{bmatrix} H & -D \\ -D' & U \end{bmatrix} \tag{39}$$

$$H = H^{-1} - H^{-1}A(A'H^{-1}A)^{-1}A'H^{-1} \tag{40a}$$

$$D = H^{-1}A(A'H^{-1}A)^{-1} \tag{40b}$$

$$U = -(A'H^{-1}A)^{-1} \tag{40c}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix}^{-1} \begin{bmatrix} g \\ b \end{bmatrix} = - \begin{bmatrix} H & -D \\ -D' & U \end{bmatrix} \begin{bmatrix} g \\ b \end{bmatrix} = \begin{bmatrix} -Hg + Db \\ D'g - Ub \end{bmatrix} \tag{41}$$

Range-Space Method / Schur-Complement Method

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix} \tag{42}$$

$$Hx - Ay = -g \quad \Leftrightarrow \quad x = H^{-1}Ay - H^{-1}g \tag{43}$$

$$b = A'x = \left(A'H^{-1}A\right) y - A'H^{-1}g \tag{44}$$

Procedure

1. Solve: $Hv = g$
2. Solve: $\left(A'H^{-1}A\right) y = b + A'v$
3. Solve: $Hx = Ay - g$

Useful when

- H is well-conditioned and easy to invert (H is diagonal or block-diagonal)
- H^{-1} is known explicitly, e.g. through quasi-Newton update formulas
- The number of equality constraints (m_a) is small, i.e. $A'H^{-1}A \in \mathbb{R}^{m_a \times m_a}$ is small.

Null-Space Method

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix} \qquad A \in \mathbb{R}^{n \times m_a} \tag{45}$$

Define the non-singular matrix $\begin{bmatrix} Y & Z \end{bmatrix} \in \mathbb{R}^{n \times n}$ with $Y \in \mathbb{R}^{n \times m_a}$ and $Z \in \mathbb{R}^{n \times (n - m_a)}$ such that

$$A' \begin{bmatrix} Y & Z \end{bmatrix} = \begin{bmatrix} A'Y & A'Z \end{bmatrix} = \begin{bmatrix} A'Y & 0 \end{bmatrix} \quad A'Y \in \mathbb{R}^{m_a \times m_a} \text{ non-singular}$$

$$x = Y x_Y + Z x_Z$$

$$-g = Hx - Ay = HYx_y + HZx_Z - Ay$$

$$b = A'x = A'Yx_y + A'Zx_Z = A'Yx_Y$$

$$\left(Z'HZ\right) x_Z = -Z'(HYx_Y + g)$$

$$\left(A'Y\right)' x_Y = Y'(Hx + g)$$

Procedure

1. Solve: $(A'Y)x_Y = b$
2. Solve: $(Z'HZ)x_Z = -Z'(HYx_Y + g)$
3. Compute: $x = Yx_Y + Zx_Z$
4. Solve: $(A'Y)'y = Y'(Hx + g)$

Null-Space Method

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix} \qquad A \in \mathbb{R}^{n \times m_a} \tag{46}$$

Define the non-singular matrix $\begin{bmatrix} Y & Z \end{bmatrix} \in \mathbb{R}^{n \times n}$ with $Y \in \mathbb{R}^{n \times m_a}$ and $Z \in \mathbb{R}^{n \times (n - m_a)}$ such that

$$A' \begin{bmatrix} Y & Z \end{bmatrix} = \begin{bmatrix} A'Y & A'Z \end{bmatrix} = \begin{bmatrix} A'Y & 0 \end{bmatrix} \quad A'Y \in \mathbb{R}^{m_a \times m_a} \text{ non-singular}$$

Procedure

1. Solve: $(A'Y)x_Y = b$
2. Solve: $(Z'HZ)x_Z = -Z'(HYx_Y + g)$
3. Compute: $x = Yx_Y + Zx_Z$
4. Solve: $(A'Y)'y = Y'(Hx + g)$

Properties

- Useful when the number of degrees of freedom, $n - m_a$, is small.
- Main drawback is its need for the null-space matrix, Z, which may be expensive to compute.

Null-Space Method: Orthonormal basis

$$A = Q \begin{bmatrix} R \\ 0 \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 R \tag{47}$$

$$A' = \begin{bmatrix} R' & 0 \end{bmatrix} Q' \tag{48}$$

$$A'Q = A' \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} = \begin{bmatrix} A'Q_1 & A'Q_2 \end{bmatrix} = \begin{bmatrix} R' & 0 \end{bmatrix} \tag{49}$$

$$\begin{bmatrix} Y & Z \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \quad A'Y = A'Q_1 = R' \tag{50}$$

Procedure

1. Solve: $R'x_Y = b$
2. Solve: $(Q_2' H Q_2) x_Z = -Q_2'(HQ_1 x_Y + g)$
3. Compute: $x = Q_1 x_Y + Q_2 x_Z$
4. Solve: $Ry = Q_1'(Hx + g)$