Quadratic Programming

- Quadratic Programming Problems
- Interior Point Methods
- Active Set Methods

Contents

Quadratic Programming (QP)

We are interested in solving quadratic programming problems of the form

$$\min_x \ \frac{1}{2}x^T H x + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

Notation:

- Hessian matrix $H \in \mathbb{R}^{n \times n}$, $H = H^T$
- gradient vector $g \in \mathbb{R}^n$
- constraint matrix $G \in \mathbb{R}^{m \times n}$
- constraint vector $b \in \mathbb{R}^m$

Quadratic Programming (QP)

We are interested in solving quadratic programming problems of the form

$$\min_x \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

Notation:

- Hessian matrix $H \in \mathbb{R}^{n\times n}$, $H = H^T$
- gradient vector $g \in \mathbb{R}^n$
- constraint matrix $G \in \mathbb{R}^{m\times n}$
- constraint vector $b \in \mathbb{R}^m$

Quadratic Programming (QP)

We are interested in solving quadratic programming problems of the form

$$\min_x \; \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b \; .$$

Notation:

- Hessian matrix $H \in \mathbb{R}^{n \times n}$, $H = H^T$
- gradient vector $g \in \mathbb{R}^n$
- constraint matrix $G \in \mathbb{R}^{m \times n}$
- constraint vector $b \in \mathbb{R}^m$

Quadratic Programming (QP)

We are interested in solving quadratic programming problems of the form

$$\min_x \ \frac{1}{2} x^T H x + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

Notation:

- Hessian matrix $H \in \mathbb{R}^{n \times n}$, $H = H^T$
- gradient vector $g \in \mathbb{R}^n$
- constraint matrix $G \in \mathbb{R}^{m \times n}$
- constraint vector $b \in \mathbb{R}^m$

Quadratic Programming (QP)

We are interested in solving quadratic programming problems of the form

$$\min_x \ \frac{1}{2} x^T H x + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

Notation:

- Hessian matrix $H \in \mathbb{R}^{n \times n}$, $H = H^T$
- gradient vector $g \in \mathbb{R}^n$
- constraint matrix $G \in \mathbb{R}^{m \times n}$
- constraint vector $b \in \mathbb{R}^m$

Some Definitions

Quadratic programming problems of the form

$$\min_x \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

are called

- feasible, if $F = \{x \mid Gx \geq b\}$ is non-empty
- bounded, if $\exists L > -\infty$ with $\frac{1}{2}x^T Hx + g^T x > L$ for all $x \in F$
- convex, if H is positive semi-definite
- strictly convex, if H is positive definite

Some Definitions

Quadratic programming problems of the form

$$\min_{x} \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

are called

- feasible, if $F = \{x \mid Gx \geq b\}$ is non-empty
- bounded, if $\exists L > -\infty$ with $\frac{1}{2}x^T Hx + g^T x > L$ for all $x \in F$
- convex, if H is positive semi-definite
- strictly convex, if H is positive definite

Some Definitions

Quadratic programming problems of the form

$$\min_x \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

are called

- feasible, if $F = \{x \mid Gx \geq b\}$ is non-empty
- bounded, if $\exists L > -\infty$ with $\frac{1}{2}x^T Hx + g^T x > L$ for all $x \in F$
- convex, if H is positive semi-definite
- strictly convex, if H is positive definite

Some Definitions

Quadratic programming problems of the form

$$\min_{x} \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

are called

- feasible, if $F = \{x \mid Gx \geq b\}$ is non-empty
- bounded, if $\exists L > -\infty$ with $\frac{1}{2}x^T Hx + g^T x > L$ for all $x \in F$
- convex, if H is positive semi-definite
- strictly convex, if H is positive definite

Sufficient Conditions for Existence of Solutions

Quadratic programming problems of the form

$$\min_x \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

have

1. a solution, if $F = \{x \mid Gx \geq b\}$ is non-empty and compact
2. no solution, if $F = \{x \mid Gx \geq b\}$ is empty
3. a unique solution, if H is strictly convex

Active and Inactive Sets

Let $\hat{x}$ be a feasible point of the QP

$$\min_x \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b\ .$$

- the set

$$\mathbb{A}(\hat{x}) = \{i \mid G_i\hat{x} = b_i\}$$

  is called the "active set" that is associated with the point $\hat{x}$.

- the set

$$\mathbb{I}(\hat{x}) = \{i \mid G_i\hat{x} > b_i\}$$

  is called the set of inactive constraints that is associated with the point $\hat{x}$.

Active and Inactive Sets

Let $\hat{x}$ be a feasible point of the QP

$$\min_x \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

- the set

  $$\mathbb{A}(\hat{x}) = \{i \mid G_i\hat{x} = b_i\}$$

  is called the "active set" that is associated with the point $\hat{x}$.

- the set

  $$\mathbb{I}(\hat{x}) = \{i \mid G_i\hat{x} > b_i\}$$

  is called the set of inactive constraints that is associated with the point $\hat{x}$.

Karush-Kuhn-Tucker Conditions

If the QP is strictly convex, there exists a unique solution x^*, an index set $\mathbb{A}$, and a multiplier y^* such that we have

1. stationarity: $Hx^* - G^T\lambda = -g$,
2. primal feasibility: $G_{\mathbb{A}}x^* = b_{\mathbb{A}}$ and $G_{\mathbb{I}}x^* \geq b_{\mathbb{I}}$
3. dual feasibility: $y^*_{\mathbb{I}} = 0$ and $y^*_{\mathbb{A}} \geq 0$.

Karush-Kuhn-Tucker Conditions

If the QP is strictly convex, there exists a unique solution x^*, an index set $\mathbb{A}$, and a multiplier y^* such that we have

1. stationarity: $Hx^* - G^T\lambda = -g$,
2. primal feasibility: $G_{\mathbb{A}}x^* = b_{\mathbb{A}}$ and $G_{\mathbb{I}}x^* \geq b_{\mathbb{I}}$
3. dual feasibility: $y^*_{\mathbb{I}} = 0$ and $y^*_{\mathbb{A}} \geq 0$.

Karush-Kuhn-Tucker Conditions

If the QP is strictly convex, there exists a unique solution x^*, an index set $\mathbb{A}$, and a multiplier y^* such that we have

1. stationarity: $Hx^* - G^T\lambda = -g$,
2. primal feasibility: $G_{\mathbb{A}}x^* = b_{\mathbb{A}}$ and $G_{\mathbb{I}}x^* \geq b_{\mathbb{I}}$
3. dual feasibility: $y^*_{\mathbb{I}} = 0$ and $y^*_{\mathbb{A}} \geq 0$.

Karush-Kuhn-Tucker Conditions

Let H be positive definite. The matrix

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix}$$

is called the KKT matrix. Important properties:

1. The KKT matrix is invertible if and only if $G_{\mathbb{A}}$ has full row-rank.
2. If $G_{\mathbb{A}}$ has full row-rank, then the multiplier y^* is unique.

Karush-Kuhn-Tucker Conditions

Let H be positive definite. The matrix

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix}$$

is called the KKT matrix. Important properties:

1. The KKT matrix is invertible if and only if $G_{\mathbb{A}}$ has full row-rank.
2. If $G_{\mathbb{A}}$ has full row-rank, then the multiplier y^* is unique.

Contents

An Equivalent Unconstrained Problem

The original QP of the form

$$\min_x \ \frac{1}{2}x^T H x + g^T x \quad \text{s.t.} \quad Gx \geq b \ .$$

can equivalently be written as

$$\min_x \ \frac{1}{2}x^T H x + g^T x + \sum_{i=1}^{m} I_{-}\left(b_i - G_i x\right) \ ,$$

where

$$I_i(z) = \left\{ \begin{array}{ll} 0 & \text{if } z \leq 0 \\ \infty & \text{otherwise} \end{array} \right\}$$

is an indicator function.

Logarithmic Barrier

The main idea of barrier method is to replace the indicator function I_{-} by a logarithmic barrier function of the form

$$L_{\mu}(z) = -\frac{1}{\mu}\log(-z)\ ,$$

where $\mu > 0$ is a parameter.

Central Path

The solution $x^*(\mu)$ of the parametric optimization problem

$$\min_x \; F(x,\mu) \quad \text{with} \quad F(x,\mu) = \frac{1}{2}x^T Hx + g^T x - \frac{1}{\mu}\sum_{i=1}^{m} \log(G_i x - b_i) \; .$$

is called the central path.

- If we have $H \succ 0$, the function F is strictly convex and smooth:

$$\begin{aligned} \nabla F(x,\mu) &= Hx + g - \frac{1}{\mu}\sum_{i=1}^{m} \frac{G_i^T}{G_i x - b_i} \\ \nabla^2 F(x,\mu) &= H + \frac{1}{\mu}\sum_{i=1}^{m} \frac{G_i G_i^T}{(G_i x - b_i)^2} \; \succ \; 0 \; . \end{aligned}$$

Central Path

The solution $x^*(\mu)$ of the parametric optimization problem

$$\min_x \; F(x,\mu) \quad \text{with} \quad F(x,\mu) = \frac{1}{2}x^T H x + g^T x - \frac{1}{\mu}\sum_{i=1}^{m} \log(G_i x - b_i)\ .$$

is called the central path.

- If we have $H \succ 0$, the function F is strictly convex and smooth:

$$\begin{aligned}
\nabla F(x,\mu) &= Hx + g - \frac{1}{\mu}\sum_{i=1}^{m} \frac{G_i^T}{G_i x - b_i} \\
\nabla^2 F(x,\mu) &= H + \frac{1}{\mu}\sum_{i=1}^{m} \frac{G_i G_i^T}{(G_i x - b_i)^2} \ \succ \ 0\ .
\end{aligned}$$

Central Path

The solution $x^*(\mu)$ can be optained by applying Newton's method for solving the optimality condition

$$\nabla F(x, \mu) \;=\; Hx + g - \frac{1}{\mu} \sum_{i=1}^{m} \frac{G_i^T}{G_i x - b_i} \;=\; 0\ .$$

If we define $\lambda_i^*(\mu) = \frac{1}{\mu(G_i x - b_i)}$, we see that $x^*(\mu)$ minimizes the Lagrangian function

$$L(x, \lambda^*(\mu)) = \frac{1}{2} x^T H x + g^T x + \sum_{i=1}^{m} \lambda_i^*(\mu)(b_i - G_i x)\ .$$

Central Path

The solution $x^*(\mu)$ can be optained by applying Newton's method for solving the optimality condition

$$\nabla F(x,\mu) \;=\; Hx + g - \frac{1}{\mu}\sum_{i=1}^{m}\frac{G_i^T}{G_i x - b_i} \;=\; 0 \; .$$

If we define $\lambda_i^*(\mu) = \frac{1}{\mu(G_i x - b_i)}$, we see that $x^*(\mu)$ minimizes the Lagrangian function

$$L(x,\lambda^*(\mu)) = \frac{1}{2}x^T H x + g^T x + \sum_{i=1}^{m}\lambda_i^*(\mu)(b_i - G_i x) \; .$$

Central Path

Now, we know from duality that

$$\begin{aligned} L(x^*(\mu), \lambda^*(\mu)) &= \frac{1}{2}(x^*(\mu))^T H x^*(\mu) + g^T(x^*(\mu)) - \frac{m}{\mu} \\ &\leq V^* \leq \frac{1}{2}(x^*(\mu))^T H x^*(\mu) \ , \end{aligned} \qquad (1)$$

where V^* is the objective value of the original QP. This analysis confirms that $x^*(\mu)$ converges to an optimal solution x^* for $\mu \to \infty$.

Barrier method

- **Input:** strictly feasible $x = x_0$, $\mu > 0$, $\rho > 1$, tolerance ϵ
- **Repeat:**

1. Solve the unconstrained optimization problem

$$\min_x \; F(x,\mu) \quad \text{with} \quad F(x,\mu) = \frac{1}{2}x^T Hx + g^T x - \frac{1}{\mu}\sum_{i=1}^{m} \log(G_i x - b_i)\ .$$

   using Newton's method "hot started" at the current iterate x.
2. Update $x = x^*(\mu)$.
3. Terminate if $\frac{m}{\mu} < \epsilon$.
4. Set $\mu \leftarrow \mu * \rho$.

Barrier method

- **Input:** strictly feasible $x = x_0$, $\mu > 0$, $\rho > 1$, tolerance ϵ
- **Repeat:**

1. Solve the unconstrained optimization problem

$$\min_x \; F(x,\mu) \quad \text{with} \quad F(x,\mu) = \frac{1}{2}x^T Hx + g^T x - \frac{1}{\mu}\sum_{i=1}^{m} \log(G_i x - b_i) \; .$$

   using Newton's method "hot started" at the current iterate x.

2. Update $x = x^*(\mu)$.
3. Terminate if $\frac{m}{\mu} < \epsilon$.
4. Set $\mu \leftarrow \mu * \rho$.

Barrier method

- **Input:** strictly feasible $x = x_0$, $\mu > 0$, $\rho > 1$, tolerance ϵ
- **Repeat:**

1. Solve the unconstrained optimization problem

$$\min_x \; F(x,\mu) \quad \text{with} \quad F(x,\mu) = \frac{1}{2}x^T Hx + g^T x - \frac{1}{\mu}\sum_{i=1}^{m} \log(G_i x - b_i) \; .$$

   using Newton's method "hot started" at the current iterate x.

2. Update $x = x^*(\mu)$.
3. Terminate if $\frac{m}{\mu} < \epsilon$.
4. Set $\mu \leftarrow \mu * \rho$.

Barrier method

- **Input:** strictly feasible $x = x_0$, $\mu > 0$, $\rho > 1$, tolerance ϵ
- **Repeat:**

1. Solve the unconstrained optimization problem

$$\min_x \; F(x,\mu) \quad \text{with} \quad F(x,\mu) = \frac{1}{2}x^T H x + g^T x - \frac{1}{\mu}\sum_{i=1}^{m} \log(G_i x - b_i) \, .$$

   using Newton's method "hot started" at the current iterate x.

2. Update $x = x^*(\mu)$.
3. Terminate if $\frac{m}{\mu} < \epsilon$.
4. Set $\mu \leftarrow \mu * \rho$.

Contents

QPs with known active set

If we would know in advance, which set $\mathbb{A}$ of constraint indices corresponds to the active set at optimal solution, it would be sufficient to solve the equality constrained QP

$$\min_x \ \frac{1}{2} x^T H x + g^T x \quad \text{s.t.} \quad G_{\mathbb{A}} x = b_{\mathbb{A}} \ .$$

If H is positive definite and $G_{\mathbb{A}}$ has full rank, this is equivalent to solving the (invertible) linear equation system

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix} \begin{pmatrix} x^* \\ -y_{\mathbb{A}}^* \end{pmatrix} = \begin{pmatrix} -g \\ b_{\mathbb{A}} \end{pmatrix}$$

QPs with known active set

If we would know in advance, which set $\mathbb{A}$ of constraint indices corresponds to the active set at optimal solution, it would be sufficient to solve the equality constrained QP

$$\min_x \ \frac{1}{2}x^T Hx + g^T x \quad \text{s.t.} \quad G_{\mathbb{A}} x = b_{\mathbb{A}} \ .$$

If H is positive definite and $G_{\mathbb{A}}$ has full rank, this is equivalent to solving the (invertible) linear equation system

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix} \begin{pmatrix} x^* \\ -y_{\mathbb{A}}^* \end{pmatrix} = \begin{pmatrix} -g \\ b_{\mathbb{A}} \end{pmatrix}$$

Primal Active Set Methods

We assume $H \succ 0$. Primal active set method start with a feasibile initial guess $x = x_0$ and an associated working set $\mathbb{A}$ and solve the equation

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ -y_{\mathbb{A}}^* \end{pmatrix} = \begin{pmatrix} -Hx - g \\ 0 \end{pmatrix}$$

The variable x is then updated by adjusting the line search parameter τ such that the next iterate

$$x^+ = x + \tau \Delta x$$

is feasible.

Primal Active Set Methods

We assume $H \succ 0$. Primal active set method start with a feasibile initial guess $x = x_0$ and an associated working set $\mathbb{A}$ and solve the equation

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ -y_{\mathbb{A}}^* \end{pmatrix} = \begin{pmatrix} -Hx - g \\ 0 \end{pmatrix}$$

The variable x is then updated by adjusting the line search parameter τ such that the next iterate

$$x^+ = x + \tau \Delta x$$

is feasible.

Blocking Constraints

In order to determine the maximum possible step length we solve

$$\max_{\tau \in [0,1]} \ \tau \quad \text{s.t.} \quad G_{\mathbb{I}}(x + \tau \Delta x) \geq b_{\mathbb{I}}$$

If we have $\tau < 1$ one of the constraint indices in $j \in \mathbb{I}$ causes a restriction on τ. In this case we update $\mathbb{A} \leftarrow A \cup \{j\}$, i.e., we add the so called "blocking constraint" to the working set.

Blocking Constraints

In order to determine the maximum possible step length we solve

$$\max_{\tau \in [0,1]} \tau \quad \text{s.t.} \quad G_{\mathbb{I}}(x + \tau \Delta x) \geq b_{\mathbb{I}}$$

If we have $\tau < 1$ one of the constraint indices in $j \in \mathbb{I}$ causes a restriction on τ. In this case we update $\mathbb{A} \leftarrow A \cup \{j\}$, i.e., we add the so called "blocking constraint" to the working set.

Removing Constraints

In another situation, we may have $\Delta x = 0$. If additionally all components of $y^*_{\mathbb{A}}$ are positive, we have found an optimal solutions. Otherwise, we drop one of the constraints that correspond to a negative component of $y^*_{\mathbb{A}}$ and determine a new step direction.

Summary: Primal Active Set Methods

Start with a feasible inital guess x_0 and working set $\mathbb{A}$ and repeat

1. Determine a step direction by solving the linear equation

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ -y_{\mathbb{A}}^* \end{pmatrix} = \begin{pmatrix} -Hx - g \\ 0 \end{pmatrix}$$

2. If $\Delta x = 0$, there are two cases possible
   - if we have $y_{\mathbb{A}}^* \geq 0$, we have found the optimal solution, terminate.
   - otherwise, update $\mathbb{A} = \mathbb{A} \setminus \{j\}$ with $(y_{\mathbb{A}}^*)_j < 0$.
3. Compute a maximum step length by solving

$$\max_{\tau \in [0,1]} \tau \quad \text{s.t.} \quad G_{\mathbb{I}}(x + \tau \Delta x) \geq b_{\mathbb{I}} \ .$$

   If $\tau < 1$ add a blocking constraint, $\mathbb{A} = \mathbb{A} \cup \{j\}$.

Summary: Primal Active Set Methods

Start with a feasible inital guess x_0 and working set $\mathbb{A}$ and repeat

1. Determine a step direction by solving the linear equation

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ -y_{\mathbb{A}}^* \end{pmatrix} = \begin{pmatrix} -Hx - g \\ 0 \end{pmatrix}$$

2. If $\Delta x = 0$, there are two cases possible
   - if we have $y_{\mathbb{A}}^* \geq 0$, we have found the optimal solution, terminate.
   - otherwise, update $\mathbb{A} = \mathbb{A} \setminus \{j\}$ with $(y_{\mathbb{A}}^*)_j < 0$.

3. Compute a maximum step length by solving

$$\max_{\tau \in [0,1]} \tau \quad \text{s.t.} \quad G_{\mathbb{I}}(x + \tau \Delta x) \geq b_{\mathbb{I}} \,.$$

   If $\tau < 1$ add a blocking constraint, $\mathbb{A} = \mathbb{A} \cup \{j\}$.

Summary: Primal Active Set Methods

Start with a feasible inital guess x_0 and working set $\mathbb{A}$ and repeat

1. Determine a step direction by solving the linear equation

$$\begin{pmatrix} H & G_{\mathbb{A}}^T \\ G_{\mathbb{A}} & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ -y_{\mathbb{A}}^* \end{pmatrix} = \begin{pmatrix} -Hx - g \\ 0 \end{pmatrix}$$

2. If $\Delta x = 0$, there are two cases possible
   - if we have $y_{\mathbb{A}}^* \geq 0$, we have found the optimal solution, terminate.
   - otherwise, update $\mathbb{A} = \mathbb{A} \setminus \{j\}$ with $(y_{\mathbb{A}}^*)_j < 0$.

3. Compute a maximum step length by solving

$$\max_{\tau \in [0,1]} \tau \quad \text{s.t.} \quad G_{\mathbb{I}}(x + \tau \Delta x) \geq b_{\mathbb{I}} \; .$$

   If $\tau < 1$ add a blocking constraint, $\mathbb{A} = \mathbb{A} \cup \{j\}$.

Dual Active Set Methods

Primal active set methods have the disadvantage that a feasible initial guess is needed. One solution to this problem is to first solve an auxiliary problem to find a feasible guess, but this is expensive in general.

An alternative are so-called dual active set methods, which apply a primal active set method to solve the dual QP

$$\max_y \; -\frac{1}{2}(G^T y - g)^T H^{-1} (G^T y - g) + y^T b \quad \text{s.t.} \quad y \geq 0 \; .$$

Here, any start point $y_0 \geq 0$ is feasible.

Dual Active Set Methods

Primal active set methods have the disadvantage that a feasible initial guess is needed. One solution to this problem is to first solve an auxiliary problem to find a feasible guess, but this is expensive in general.

An alternative are so-called dual active set methods, which apply a primal active set method to solve the dual QP

$$\max_{y} \; -\frac{1}{2}(G^T y - g)^T H^{-1}(G^T y - g) + y^T b \quad \text{s.t.} \quad y \geq 0 \; .$$

Here, any start point $y_0 \geq 0$ is feasible.