16 Quadratic Programming
Quadratic programming

- Quadratic object function + linear constraints

\[
\min_x q(x) = \frac{1}{2} x^T G x + x^T c \\
\text{s.t.} \quad a_i^T x = b_i \quad i \in \mathbf{E} \\
\quad a_i^T x \geq b_i \quad i \in \mathbf{I}
\]

- If $G$ is positive semidefinite, it is called *convex* QP.
- If $G$ is positive definite, it is called *strictly convex* QP.
- If $G$ is indefinite, it is called *nonconvex* QP.
Equality constraints

- First order condition for optimality (chap 12)
  
  \[
  \begin{pmatrix}
  G & A^T \\
  A & 0
  \end{pmatrix}
  \begin{pmatrix}
  -p \\
  \lambda^*
  \end{pmatrix}
  =
  \begin{pmatrix}
  g \\
  h
  \end{pmatrix}
  \]

- where \( h = Ax - b \), \( g = c + Gx \), and \( p = x^* - x \). Pair \((x^*, \lambda^*)\) is the optimal solution and Lagrangian multipliers.

- \( K = \begin{pmatrix}
  G & A^T \\
  A & 0
  \end{pmatrix} \) is called the KKT matrix, which is indefinite, nonsingular if the project Hessian \( Z^T GZ \) is positive definite. (chap 12)
Solving the KKT system

- Direct methods
  - Symmetric indefinite factorization $P^T KP = LBL^T$.
  - Schur complement method (1)
  - Null-Space method (2)

- Iterative methods
  - CG applied to the reduced system (3)
  - The projected CG method (4)
  - Newton-Krylov method
1. Schur-complement method

- Assume G is invertible.

\[
\begin{pmatrix}
I & 0 \\
AG^{-1} & -I
\end{pmatrix}
\begin{pmatrix}
G & A^T \\
A & 0
\end{pmatrix}
\begin{pmatrix}
-p \\
\lambda^*
\end{pmatrix}
= 
\begin{pmatrix}
I & 0 \\
AG^{-1} & -I
\end{pmatrix}
\begin{pmatrix}
g \\
h
\end{pmatrix}
\]

\[
\begin{pmatrix}
G & A^T \\
0 & AG^{-1}A^T
\end{pmatrix}
\begin{pmatrix}
-p \\
\lambda^*
\end{pmatrix}
= 
\begin{pmatrix}
g \\
AG^{-1}g - h
\end{pmatrix}
\]

- This gives two equations

  - Solve for \( \lambda^* \) first.
  - Then use \( \lambda^* \) to solve \( p \).

\[-Gp + A^T \lambda^* = g\]
\[(AG^{-1}A^T)\lambda^* = AG^{-1}g - h\]

- Matrix \( AG^{-1}A^T \) is called the Schur-complement
2. Null space method

- Require (1) $A$ has full row rank (2) $Z^TGZ$ is positive definite. ($G$ itself can be singular.)

- QR decomposition of $A^T$, and partition $p$ ($Q_2 = Z$

$$A^T = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R \\ 0 \end{pmatrix}, p = Q_1p_1 + Q_2p_2$$

$$\begin{pmatrix} G & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} -p \\ \lambda^* \end{pmatrix} = \begin{pmatrix} G & Q_1R \\ R^TQ_1^T & 0 \end{pmatrix} \begin{pmatrix} -p \\ \lambda^* \end{pmatrix} = \begin{pmatrix} g \\ h \end{pmatrix}$$

$$Rp_1 = -h$$

$$(Q_2^T G Q_2)p_2 = -Q_2^T G Q_1p_1 - Q_2^T g$$

$$R^T \lambda^* = Q_1^T (g + Gp)$$
3. Conjugate Gradient for $Z^T G Z$

- Assume $A^T = Q_1 R$, $Q_2 = Z$. (see previous slide).
  - The solution can be expressed as $x = Q_1 x_1 + Q_2 x_2$.
  - $x_1 = R^{-T} b$ cannot be changed. Reduce the original problem to the unconstrained one (see chap 15)
    $$\min_{x_2} \frac{1}{2} x_2^T Q_2^T G Q_2 x_2 + x_2^T Q_2^T c$$
- The minimizer is the solution of (see chap 2)
  \[ Q_2^T G Q_2 x_2 = -Q_2^T c \]
  - If $Q_2^T G Q_2$ is s.p.d., the CG can be used as a solver. (see chap 5)
4. The Projected CG method

- Use the same assumptions as the previous slide
- Define $P = Q_2 Q_2^T$, an orthogonal projector, by which $Px \in \text{span}(Q_2)$ for all $x$.
- The projected CG method uses projected residual $r_{k+1}^p = Pr_{k+1}$ (see chap 5) instead of $r_{k+1}$.
- Matrix $P$ is equivalent to $P = I - A^T (AA^T)^{-1} A$.
  - Numerically unstable, use iterative refinement to improve the accuracy.
Inequality constraints

- Review the optimality conditions (chap 12)
  \[ Gx^* + c - \sum_{i \in A(x^*)} \lambda_i^* a_i = 0 \]
  \[ a_i^T x^* = b_i \quad i \in E \]
  \[ a_i^T x^* \geq b_i \quad i \in I \setminus A(x^*) \]
  \[ \lambda^* \geq 0 \quad i \in I \cap A(x^*) \]

- Two difficulties (Figure 16.1, 16.2)
  - Nonconvexity: more than one minimizer
  - Degeneracy: The gradients of active constraints are linearly dependent
Methods for QP

1. Active set method for convex QP
   ● Only for small or medium sized problems

2. Interior point method
   ● Good for large problems

3. Gradient projection method
   ● Good when the constraints are bounds for variables
1. Active set method for QP

- Basic strategy
  1. Find an active set (called working set, denoted \( W_k \))
  2. Check if \( x_k \) is optimal
  3. If not, find \( p_k \) (solving an equality constrained QP)
     \[
     \min_{p} \frac{1}{2} p^T G p + g^T p
     \]
     s.t. \( a_i^T p = 0 \quad i \in W_k \)
  4. Find \( \alpha_k \in [0,1] \) such that \( x_{k+1} = x_k + \alpha_k p_k \) satisfies all constraints \( a_i, i \notin W_k \) (next slide)
  5. If \( \alpha_k < 1 \), update \( W_k \).
Step length

- Need find $\alpha_k \in [0, 1]$ as large as possible such that $a_i^T(x_k + \alpha_k p_k) \geq b_i, \ i \not\in W_k$
  - If $a_i^T p_k \geq 0$, then for all $\alpha_k \geq 0$, it satisfies anyway
  - If $a_i^T p_k < 0$, $\alpha_k$ need be $\leq (b_i - a_i^T x_k)/a_i^T p_k$.

$$\alpha_k = \min \left( 1, \ \min_{i \not\in W_k, a_i^T p_k < 0} \frac{b_i - a_i^T x_k}{a_i^T p_k} \right)$$
2. Interior point method for QP

- Please review the IPM for LP (chap 14)
- The problem \( \min_x q(x) = \frac{1}{2} x^T G x + x^T c \)
  s.t. \( Ax \geq b \)
- KKT conditions
  \[
  G x - A^T \lambda + c = 0 \\
  A x - y - b = 0 \\
  y_i \lambda_i = 0, \ i = 1, 2, \ldots, m \\
  y, \lambda \geq 0
  \]
- Complementary measure:
  \[
  \mu = y^T \lambda / m
  \]
Nonlinear equation

\[ F(x, y, \lambda, \sigma \mu) = \begin{pmatrix} Gx - A^T \lambda + c \\ Ax - y - b \\ Y \Lambda e - \sigma \mu e \end{pmatrix} \]

Using Newton’s method to solve \( F(x, y, \lambda, \sigma \mu) = 0 \)

\[
\begin{pmatrix} G & 0 & -A^T \\ A & -I & 0 \\ 0 & \Lambda & Y \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta \lambda \end{pmatrix} = \begin{pmatrix} -r_d \\ -r_p \\ -\Lambda Ye + \sigma \mu e \end{pmatrix}
\]

where \( r_d = Gx - A^T \lambda + c \)
\( r_p = Ax - y - b \)

Next step

\[
(x_{k+1}, y_{k+1}, \lambda_{k+1}) = (x_k, y_k, \lambda_k) + \alpha_k(\Delta x, \Delta y, \Delta \lambda)
\]
3. Gradient projection method

- Consider the bounded constrained problem
  \[ \min_x q(x) = \frac{1}{2} x^T G x + x^T c \]
  s.t. \[ l \leq x \leq u, \]
  - The gradient of \( q(x) \) is \( g = Gx + c \).

- Bounded projection
  \[ P(x, l, u) = \begin{cases} 
  l & \text{if } x < l \\
  x & \text{if } l < x < u \\
  u & \text{if } x > u 
  \end{cases} \]
  - The gradient projection is a piecewise linear path
  \[ x(t) = P(x-tg, l, u) \]
**Piecewise linear path**

\[ x(t) = \begin{cases} 
  x_1 + \Delta t p_1 & \text{for } 0 \leq \Delta t < t_1 \quad (x_1 = x) \\
  x_2 + \Delta t p_2 & \text{for } t_1 \leq \Delta t < t_2 \\
  \vdots \\
  x_n + \Delta t p_n & \text{for } t_{n-1} \leq \Delta t \leq t 
\end{cases} (x_n = x_{n-1} + t_{n-1} p_{n-1}) \]

**For each dimension**

\[ t_i = \begin{cases} 
  (x_i - u_i) / g_i & \text{if } g_i < 0 \\
  (x_i - l_i) / g_i & \text{if } g_i > 0 \\
  \infty & \text{otherwise} 
\end{cases} \]

\[ x_i(t) = \begin{cases} 
  x_i - t g_i & \text{if } t \leq t_i \\
  x_i - t_i g_i & \text{otherwise} 
\end{cases} \]

\[ p_i(t) = \begin{cases} 
  -g_i & \text{if } t_{i-1} \leq t_i \\
  0 & \text{otherwise} 
\end{cases} \]