Chapter 12: Quadratic Programming

Outline

Outline

Introduction

- An optimization problem with a quadratic objective function and linear constraints is called a *quadratic program*.
- Problems of this type are important in their own right, and they also arise as subproblems in methods for general constrained optimization, such as sequential quadratic programming , augmented Lagrangian methods, and interior-point methods.
- The general quadratic program (QP) can be stated as

$$\min_{x\in\mathbb{R}^n} \quad q(x) = \frac{1}{2}x^T Gx + x^T c \tag{1.1a}$$

$$s.t. \quad a_i^T x = b_i, \qquad i \in \mathcal{E}, \tag{1.1b}$$

$$a_i^T x \geq b_i, \qquad i \in \mathcal{I}, \tag{1.1c}$$

where G is a symmetric $n \times n$ matrix, $\mathcal{E}$ and $\mathcal{I}$ are finite sets of indices, and d, x, and $\{a_i\}$, $i \in \mathcal{E} \cup \mathcal{I}$, are vectors with n elements.

Introduction

Quadratic programs can always be solved (or can be shown to be infeasible) in a finite number of iterations, but the effort required to find a solution depends strongly on the characteristics of the objective function and the number of inequality constraints.

- If the Hessian matrix G is positive semidefinite, we say that (1.1) is a convex QP, and in this case the problem is sometimes not much more difficult to solve than a linear program.
- Nonconvex QPs, in which G is an indefinite matrix, can be more challenging, since they can have several stationary points and local minima.

We focus on studying algorithms that find the solution of a convex quadratic program or a stationary point of a general (nonconvex) quadratic program.

Example: Portfolio Optimization

Try to find a portfolio for which the expected return is large while the risk is small. Define the model:

$$\begin{aligned} \max_x \quad & x^T\mu - \kappa x^T G x \\ s.t. \quad & \sum_{i=1}^n x_i = 1, \\ & x \geq 0. \end{aligned}$$

Outline

Properties of Equality-Constrained QPs

Equality-constrained QP:

$$\begin{aligned} \min \quad & q(x) \equiv \frac{1}{2}x^T Gx + x^T c && \text{(2.1a)} \\ s.t. \quad & Ax = b, && \text{(2.1b)} \end{aligned}$$

where A is the $m \times n$ Jacobian of constraints (with $m \leq n$) whose rows are a_i^T, $i \in \mathcal{E}$ and b is the vector in $\mathbb{R}^m$ whose components are b_i, $i \in \mathcal{E}$. For the present, we assume that A has full row rank (rank m) so that the constraints are consistent.

Properties of Equality-Constrained QPs

The first-order necessary conditions for x^* to be a solution of (2.1) state that there is a vector λ^* such that the following system of equations is satisfied:

$$\begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}.$$

This system can be rewritten in a form that is useful for computation by expressing x^* as

$$x^* = x + p,$$

where x is some estimate of the solution and p is the desired step. By introducing this notation and rearranging the equations, we obtain

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -p \\ \lambda^* \end{bmatrix} = \begin{bmatrix} g \\ h \end{bmatrix}, \tag{2.2}$$

where

$$h = Ax - b, \qquad g = c + Gx, \qquad p = x^* - x.$$

Properties of Equality-Constrained QPs

The matrix in (2.2) is called the Karush-Kuhn-Tucker (KKT) matrix. We use Z to denote the $n \times (n-m)$ matrix whose columns are a basis for the null space of A. That is, Z has full rank and satisfies $AZ = 0$.

Theorem 1

Let A have full row rank, and assume that the reduced-Hessian matrix $Z^T GZ$ is positive definite. Then the KKT matrix

$$K = \begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \tag{2.3}$$

is nonsingular, and there is a unique vector pair (x^*, λ^*) satisfying (2.2). Furthermore, x^* is the unique global solution of (2.1).

Proof idea: for any feasible point x (with $Ax = b$),

$$q(x) = \frac{1}{2} p^T Gp + q(x^*), \quad \text{where } p = x^* - x.$$

Properties of Equality-Constrained QPs

- When the reduced Hessian matrix $Z^T G Z$ is positive semidefinite with zero eigenvalues, the vector x^* satisfying (2.2) is a local minimizer but not a strict local minimizer.
- If the reduced Hessian has negative eigenvalues, then x^* is only a stationary point, but not a local minimizer.

Direct Solution of the KKT System

In this subsection we discuss efficient methods for solving the KKT system (2.2), namely

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -p \\ \lambda^* \end{bmatrix} = \begin{bmatrix} g \\ h \end{bmatrix}.$$

The first important observation is that if $m \geq 1$, the KKT matrix is always indefinite.

Theorem 2

Suppose that A has full row rank and that the reduced Hessian $Z^T G Z$ is positive definite. Then the KKT matrix (2.3) has n positive eigenvalues, m negative eigenvalues, and no zero eigenvalues.

Factoring the Full KKT System

- One option for solving (2.2) is to perform a triangular factorization on the full KKT matrix and then perform backward and forward substitution with the triangular factors.
- We cannot use the Cholesky factorization algorithm because the KKT matrix is indefinite.
- We could use Gaussian elimination (LU factorization), but it ignores the symmetry.
- Currently, the most effective strategy in this case is to use a *symmetric indefinite factorization*, namely

$$P^T KP = LBL^T.$$

  This approach of factoring the full $(n+m) \times (n+m)$ KKT matrix (2.3) is quite effective on some problems.
- It may be expensive, however, when the heuristics for choosing the permutation matrix P are not able to maintain sparsity in the L factor, so that L becomes much more dense than the original coefficient matrix.

Schur-Complement Method

Assuming that G is positive definite, we can multiply the first equation in (2.2) by AG^{-1} and then subtract the second equation to obtain a linear system in the vector λ^* alone:

$$(AG^{-1}A^T)\lambda^* = (AG^{-1}g - h). \tag{2.4}$$

We solve this symmetric positive definite system for λ^* and then recover p from the first equation in (2.2) by solving

$$Gp = A^T\lambda^* - g. \tag{2.5}$$

This approach requires us to perform operation with G^{-1}, as well as to compute the factorization of the $m \times m$ matrix $AG^{-1}A^T$. Therefore, it is most useful when:

- G is well conditioned and easy to invert;
- G^{-1} is known explicitly through a quasi-Newton updating formula; or
- the number of equality constraints m is small, so that the number of backsolves needed to form the matrix $AG^{-1}A^T$ is not too large.

Null-Space Method

- The null-space method does not require nonsingularity of G and therefore has wider applicability.
- It assumes only that A have full row rank, and assume that the reduced-Hessian matrix $Z^T GZ$ is positive definite.
- However, it requires knowledge of the null-space basis matrix Z. Like the Schur-complement method, it exploits the block structure in the KKT system to decouple (2.3) into smaller systems.
- Suppose $Z \in \mathbb{R}^{n\times(n-m)}$ is a basis matrix of Null(A) and $Y \in \mathbb{R}^{n\times m}$ such that $[Y|Z]$ is nonsingular. Then the vector p in (2.2) can be expressed as

$$p = Yp_Y + Zp_Z, \tag{2.6}$$

where $p_Y \in \mathbb{R}^m$, $p_Z \in \mathbb{R}^{n-m}$.

Null-Space Method

- By substituting p into the second equation of (2.2) and recalling that $AZ = 0$, we obtain
$$(AY)p_Y = -h.$$
- Since A has rank m and $[Y|Z]$ is $n \times n$ nonsingular, the product $A[Y|Z] = [AY|0]$ has rank m. Therefore, AY is a nonsingular $m \times m$ matrix, and p_Y is well determined by the above equations.
- Meanwhile, we can substitute (2.6) into the first equation of (2.2) to obtain p_Z by solving
$$(Z^T GZ)p_Z = -Z^T GYp_Y - Z^T g. \tag{2.7}$$
- To obtain the Lagrange multiplier, we multiply the first equation of (2.2) by Y^T to obtain the linear system
$$(AY)^T \lambda^* = Y^T(g + Gp),$$

  which can be solved for λ^*.

Null-Space Method

- The null-space approach can be effective when the number of degrees of freedom $n - m$ is small.
- Its main drawback is the need for the null-space matrix Z, which can be expensive to compute in many large problems.
- The matrix Z is not uniquely defined, and if it is poorly chosen, the reduced system (2.7) may become ill conditioned.
- If we choose Z to have orthonormal columns, as is normally done in software for small and medium-sized problems, then the conditioning of $Z^T G Z$ is at least as good as that of G itself.
- When A is large and sparse, however, this choice of Z is relatively expensive to compute, so for practical reasons we are often forced to use one of the less reliable choices of Z.

Iterative Solution of the KKT System

An alternative to the direct factorization techniques is to use an iterative method to solve the KKT system (2.2). Iterative methods are suitable for solving very large systems and often lend themselves well to parallelization.

- The CG method is not recommended for solving the full system (2.2), because it can be unstable on systems that are not positive definite.
- Better options are Krylov methods for general linear or symmetric indefinite systems. Candidates include the GMRES, QMR, and LSQR methods.
- Other iterative methods can be derived from the null-space approach by applying the CG method to the reduced system (2.7), namely

$$(Z^T GZ)P_Z = -Z^T Gp_Y - Z^T g.$$

Outline

Inequality-Constrained Quadratic Programs

We discuss several classes of algorithms for solving QPs that contain inequality constraints and possibly equality constraints.

- *Classical active-set methods* can be applied both to convex and nonconvex problems, and they have been the most widely used methods since the 1970s.
- *Gradient-projection methods* attempt to accelerate the solution process by allowing rapid changes in the active set, and are most efficient when the only constraints in the problem are bounds on the variables.
- *Interior-point methods* have recently been shown to be effective for solving large convex quadratic programs.

Optimality Conditions for Inequality-Constrained Problems

The Lagrangian for (1.1) is

$$\mathcal{L}(x,\lambda) = \frac{1}{2}x^T Gx + x^T c - \sum_{i\in\mathcal{I}\cup\mathcal{E}} \lambda_i(a_i^T x - b_i). \tag{3.1}$$

The active set $\mathcal{A}(x^*)$ consists of the indices of the constraints for which equality holds at x^*:

$$\mathcal{A}(x^*) = \{i \in \mathcal{E}\cup\mathcal{I} | a_i^T x^* = b_i\}. \tag{3.2}$$

Optimality Conditions for Inequality-Constrained Problems

By specializing the KKT conditions to QP (1.1), we find that any solution x^* satisfies the following first-order conditions, for some Lagrangian multiplier λ_i^*, $i \in \mathcal{A}(x^*)$:

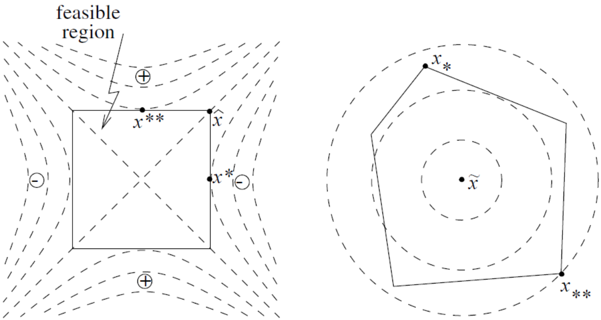
$$
\begin{aligned}
Gx^* + c - \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* a_i &= 0, && && (3.3a)\\
a_i^T x^* &= b_i, && \forall i \in \mathcal{A}(x^*), && (3.3b)\\
a_i^T x^* &\geq b_i, && \forall i \in \mathcal{I} \backslash \mathcal{A}(x^*), && (3.3c)\\
\lambda_i^* &\geq 0, && \forall i \in \mathcal{I} \cap \mathcal{A}(x^*). && (3.3d)
\end{aligned}
$$

Theorem 3

If x^* satisfies the conditions (3.3) for some λ_i^*, $i \in \mathcal{A}(x^*)$, and G is positive semidefinite, then x^* is global solution of (1.1). When G is positive definite, x^* is actually the unique global solution.

Optimality Conditions for Inequality-Constrained Problems

When G is not positive definite, the general problem (1.1) may have more than one strict local minimizer at which the second-order necessary conditions are satisfied. Such problems are referred to as being "nonconvex" or "indefinite", and they cause some complication for algorithms. Examples of indefinite quadratic programs are illustrated in the following figure.



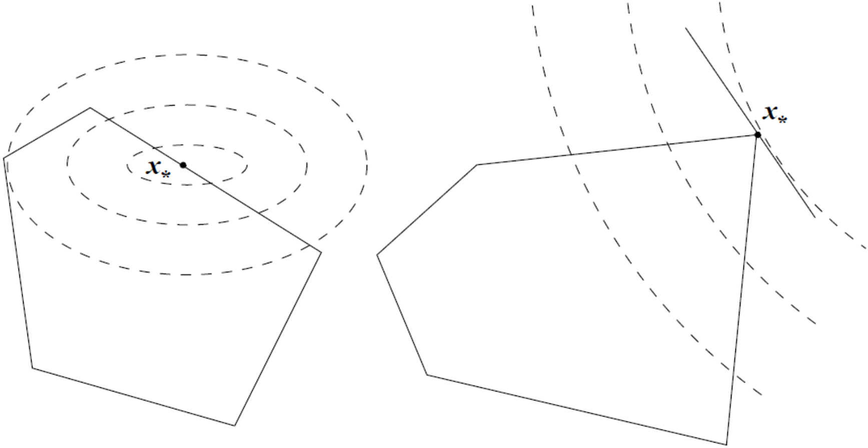
Degeneracy

A second property that causes difficulties for some algorithms is degeneracy. Unfortunately, this term has been given a variety of meanings, which can cause confusion. Essentially, it refers to situations in which either

(a) the active constraint gradients a_i, $i \in \mathcal{A}(x^*)$, are linearly dependent at the solution x^*, and/or

(b) the strict complementarity condition fails to hold, that is, there is some index $i \in \mathcal{A}(x^*)$ such that all Lagrange multipliers satisfying (3.3) have $\lambda^* = 0$.

Degeneracy



Degeneracy

Degeneracy can cause problems for algorithms for two main reasons.

- First, linear independence of the active constraint gradients can cause numerical difficulties in the step computation because certain matrices that we need to factor become rank deficient.
- Second, when the problem contains weakly active constraints, it is difficult for the algorithm to determine whether these constraints are active at the solution. In the case of active-set methods and gradient projection methods, this indecisiveness can cause the algorithm to zigzag as the iterates move on and off the weakly active constraints on successive iterations. Safeguards must be used to prevent such behavior.

Active-Set Methods for Convex QPs

- We now describe active-set methods, which are generally the most effective methods for small- to medium-scale problems.
- We consider only the convex case, in which the matrix G is positive semidefinite. Since the feasible region is a convex set, any local solution of the QP is a global minimizer.
- The case in which G is an indefinite matrix raises complications in the algorithms and is outside the scope of in our class. For a discussion of nonconvex QPs, we refer to

N. I. M. Gould, D. Orban, and P.L. Toint, *Numerical methods for large-scale nonlinear optimization*, Acta Numerica, 14(2005), pp. 299-361.

Active-Set Methods for Convex QPs

If the contents of the *optimal* active sets were known in advance, we could find the solution x^* by applying one of the techniques for equality-constrained QP to the problem

$$\begin{aligned} \min_x \quad & q(x) = \frac{1}{2} x^T G x + x^T c && \text{(3.4a)} \\ s.t. \quad & a_i^T x = b_i, \;\; i \in \mathcal{A}(x^*). && \text{(3.4b)} \end{aligned}$$

Active-set methods for QP differ from the simplex method, which is an active-set method for LP, in that the iterates (and the solution x^*) are not necessarily vertices of the feasible region.

Active-Set Methods for Convex QPs

Active-set methods for QP come in three varieties, known as

primal, *dual*, and *primal-dual*.

We restrict our discussion to primal methods, which generate iterates that remain feasible with respect to the primal problem (1.1) while steadily decreasing the primal objective function $q(\cdot)$.

Active-Set Methods for Convex QPs

- Primal active-set methods usually start by computing a feasible initial iterate x_0, and then ensure that all subsequent iterates remain feasible.
- They find a step from one iterate to the next by solving an equality constrained quadratic subproblem in which a subset of the constraints is imposed as equalities.
- This subset is referred to as the *working set* and is denoted at the k-th iterate x_k by $\mathcal{W}_k$. It consists of all the equality constraints $i \in \mathcal{E}$ together with some - but not necessarily all - of the active inequality constraints at x_k, namely $\mathcal{E} \subseteq \mathcal{W}_k \subseteq \mathcal{A}(x_k)$.
- An important requirement we impose on $\mathcal{W}_k$ is that the gradients a_i of the constraints in the working set be linearly independent, even when the full set of active constraints at that point has linearly dependent gradients.

Active-Set Methods for Convex QPs

- Given an iterate x_k and the working set $\mathcal{W}_k$, we first check whether x_k minimizes the quadratic q in the subspace defined by the working set.
- If not, we compute a step p by solving the equality-constrained QP subproblem

$$\min_{p} \quad \frac{1}{2}p^T G p + g_k^T p \tag{3.5a}$$

$$s.t. \qquad a_i^T p = 0, \qquad i \in \mathcal{W}_k \tag{3.5b}$$

with its solution denoted by p_k.

Active-Set Methods for Convex QPs

- Suppose for the moment that the optimal p_k from (3.5) is nonzero.
- If $x_k + p_k$ is feasible with respect to all the constraints, we set $x_{k+1} = x_k + p_k$. Otherwise, we set

$$x_{k+1} = x_k + \alpha_k p_k,$$

where the step-length parameter α_k is chosen to be the largest value in the range $[0, 1)$ for which all constraints are satisfied.

- If $a_i^T p_k \geq 0$ for some $i \notin \mathcal{W}_k$, then for all $\alpha_k \geq 0$ we have $a_i^T(x_k + \alpha_k p_k) \geq a_i^T x_k \geq b_i$. Hence, this constraint will be satisfied for all nonnegative choices of the step-length parameter.
- If $a_i^T p_k < 0$ for some $i \notin \mathcal{W}_k$, however, we have that $a_i^T(x_k + \alpha_k p_k) \geq b_i$ only if

$$\alpha_k \leq \frac{b_i - a_i^T x_k}{a_i^T p_k}. \tag{3.6}$$

Active-Set Methods for Convex QPs

- Since we want α_k to be as large as possible in $[0, 1]$ subject to retaining feasibility, we have the following definition:

$$\alpha_k \equiv \min\left(1, \min_{i \notin \mathcal{W}_k, a_i^T p_k < 0} \frac{b_i - a_i^T x_k}{a_i^T p_k}\right). \tag{3.7}$$

- We call the constraints i for which the minimum is achieved the *blocking constraints*. (If $\alpha_k = 1$ and no new constraints are active at $x_k + \alpha_k p_k$, then there are no blocking constraints on this iteration.)
- Note that it is quite possible for α_k to be zero, since we could have $a_i^T p_k < 0$ for some constraint $i \notin \mathcal{W}$ that is active at x_k.

Active-Set Methods for Convex QPs

- If $\alpha_k < 1$, that is, the step along p_k was blocked by some constraint not in $\mathcal{W}_k$, a new working set $\mathcal{W}_{k+1}$ is constructed by adding one of the blocking constraints to $\mathcal{W}_k$.
- We continue to iterate in this manner, adding constraints to the working set until we reach a point $\hat{x}$ that minimizes the quadratic objective function over its current working set $\hat{\mathcal{W}}$.
- Then $p = 0$ is the solution of (3.5). Then there exist some Lagrange multipliers $\hat{\lambda}_i$, $i \in \hat{\mathcal{W}}$ such that

$$\sum_{i \in \hat{\mathcal{W}}} a_i \hat{\lambda}_i = g_k = G\hat{x} + c \tag{3.8}$$

- Define $\hat{\lambda}_i = 0$ for $i \notin \hat{\mathcal{W}}$. Together with the feasibility of $\hat{x}$, it follows that $\hat{x}$ and $\hat{\lambda}$ satisfy the first three conditions (3.3a)-(3.3c) in KKT conditions.

Active-Set Methods for Convex QPs

- We now examine the signs of $\hat{\lambda}_i$, $i \in \mathcal{W} \cap \mathcal{I}$.
- If $\hat{\lambda}_i \geq 0$, $i \in \hat{\mathcal{W}} \cap \mathcal{I}$, (3.3d) is also satisfied, so we conclude that $\hat{x}$ is a KKT point. Since G is positive semidefinite, $\hat{x}$ is a local minimizer. When G is positive definite, $\hat{x}$ is a strict local minimizer.
- If $\exists\, j \in \hat{\mathcal{W}} \cap \mathcal{I}$ such that $\hat{\lambda}_j < 0$, we can remove j from $\hat{\mathcal{W}}$ and solve a new subproblem (3.5) for the new step. It can be shown that this strategy produces a direction p at the next iteration that is feasible with respect to the dropped constraint.
- While any index j for which $\hat{\lambda}_j < 0$ usually will give directions along which the algorithm can make progress, the most negative multiplier is often chosen in practice.

Active-Set Methods for Convex QPs

Theorem 4

Suppose that the solution p_k of (3.5) is nonzero and satisfies the second-order sufficient conditions for optimality for that problem. Then the function $q(\cdot)$ is strictly decreasing along the direction p_k.

Active-Set Method for Convex QP

Compute a feasible starting point x_0;
Set $\mathcal{W}_0$ to be a subset of the active constraints at x_0;
for $k = 0, 1, 2, \ldots$
  Solve (16) to find p_k;
  if $p_k = 0$
    Compute Lagrange multipliers $\hat{\lambda}_i$ that satisfy (18),
      set $\hat{\mathcal{W}} = \mathcal{W}_k$;
    if $\hat{\lambda}_i \geq 0$ for all $i \in \mathcal{W}_k \cap \mathcal{I}$;
      STOP with solution $x^* = x_k$;
    else
      Set $j = \arg\min_{j \in \mathcal{W}_k \cap \mathcal{I}} \hat{\lambda}_j$;
      $x_{k+1} = x_k$; $\mathcal{W}_{k+1} \leftarrow \mathcal{W}_k \backslash \{j\}$;
  else (* $p_k \neq 0$ *)
    Compute α_k from (17);
    $x_{k+1} \leftarrow x_k + \alpha_k p_k$;
    if there are blocking constraints
      Obtain $\mathcal{W}_{k+1}$ by adding one of the blocking
        constraints to $\mathcal{W}_{k+1}$;
    else
      $\mathcal{W}_{k+1} \leftarrow \mathcal{W}_k$;
end (for)

Further Remarks on the Active-Set Method

- Initial feasible point;
- Finite termination of active-set algorithm on strictly convex QPs;
- Updating Factorizations.

The Bound-constrained Problem

Consider the bound-constrained problem:

$$\min_{x} \quad q(x) = \frac{1}{2}x^T G x + x^T c \tag{3.9a}$$

$$s.t. \qquad l \le x \le u, \tag{3.9b}$$

where G is symmetric and l and u are vectors of lower and upper bounds on the components of x.

- The feasible region is sometimes called a "box" because of its rectangular shape.
- Some components of x may lack an upper or a lower bound; we handle these cases by setting the appropriate components of l and u to $-\infty$ and $+\infty$, respectively.
- We do not make any positive definiteness assumptions on G, since the gradient projection approach can be applied to both convex and nonconvex problems.

Optimality Conditions for (3.9)

- The KKT conditions for (3.9) is

$$
\begin{aligned}
Gx + c &= \lambda_1 - \lambda_2, && (3.10a)\\
l \leq x &\leq u, && (3.10b)\\
\lambda_1 \geq 0, \quad \lambda_2 &\geq 0, && (3.10c)\\
\lambda_1 \circ (x - l) = 0, \quad \lambda_2 \circ (x - u) &= 0. && (3.10d)
\end{aligned}
$$

- Equivalently, if x is a KKT point, then

$$
\begin{aligned}
x - P(x - (Gx + c), l, u) &= 0,\\
l \leq x &\leq u,
\end{aligned}
$$

where

$$
P(x, l, u)_i = \begin{cases} l_i, & \text{if } x_i < l_i,\\ x_i, & \text{if } x_i \in [l_i, u_i],\\ u_i, & \text{if } x_i > u_i. \end{cases} \tag{3.12}
$$

The Gradient Projection Method

Each iteration of the gradient projection algorithm consists of two stages.

- In the first stage, we search along the steepest descent direction from the current point x, that is, the direction $-g$, where $g = Gx + c$. When a bound is encountered, the search direction is "bent" so that it stays feasible. We search along this piecewise path

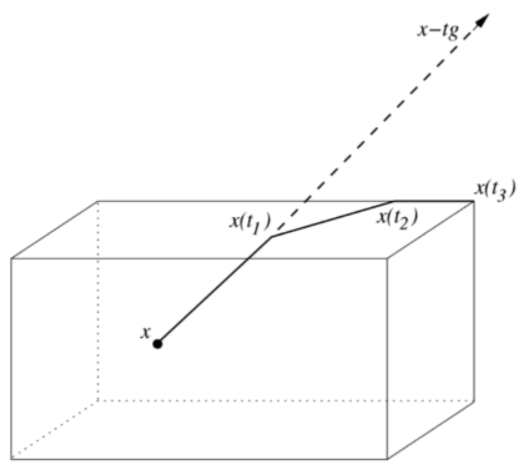
$$x(t) = P(x - tg, l, u)$$

  and locate the first local minimizer of q, which we denote by x^c and refer to as the Cauchy point. The working set is now defined to be the set of bound constraints that are active at the Cauchy point, denoted by $\mathcal{A}(x^c)$.

- In the second stage of each gradient projection iteration, we "explore" the face of the feasible box on which the Cauchy point lies by solving a subproblem in which the active components x_i for $i \in \mathcal{A}(x^c)$ are fixed at the values x_i^c. Namely, solve

$$\begin{array}{ll} \min & q(x) = \frac{1}{2}x^T Gx + x^T c \\ s.t. & x_i = x_i^c,\ i \in \mathcal{A}(x^c), \\ & l_i \leq x_i \leq u_i,\ i \notin \mathcal{A}(x^c). \end{array}$$

The Piecewise-Linear Path



Gradient Projection Method for QP

Algorithm

Compute a feasible starting point x_0;

for $k = 0, 1, 2, \ldots$

if x_k satisfies the KKT conditions for (3.9) (or $\|x_k - P(x_k - g_k, l, u)\| = 0$)

stop with solution $x_* = x_k$;

Set $x = x_k$ and find the Cauchy point x^c;

Find an approximate solution x^+ of (3.13) such that $q(x^+) \leq q(x^c)$

and x^+ is feasible;

$x_{k+1} \leftarrow x^+$;

end(for)

Interior-Point Methods

Consider convex quadratic programs with inequality constraints:

$$\min_x \quad q(x) \equiv \frac{1}{2}x^T G x + x^T c \tag{3.14a}$$

$$s.t. \qquad Ax \geq b, \tag{3.14b}$$

where G is symmetric and positive semidefinite, and where the $m \times n$ matrix A and righthand-side b are defined by

$$A = [a_i]_{i\in\mathcal{I}}, \qquad b = [b_i]_{i\in\mathcal{I}}, \qquad \mathcal{I} = \{1, 2, \cdots, m\}.$$

Interior-Point Methods

The KKT conditions for problem (3.14) can be written as

$$\begin{aligned} Gx - A^T\lambda + c &= 0, && && (3.15a)\\ Ax - y - b &= 0, && && (3.15b)\\ y_i\lambda_i &= 0, && i = 1, 2, \cdots, m, && (3.15c)\\ (y, \lambda) &\geq 0. && && (3.15d) \end{aligned}$$

When G is positive semidefinite, these KKT conditions are not only necessary but also sufficient, so we can solve the convex program (3.14) by finding solutions of the system (3.15).

Interior-Point Methods

Given a current iterate (x, y, λ) that satisfies $(y, \lambda) > 0$, we can define a complementarity measure

$$\mu = \frac{y^T \lambda}{m}.$$

The path-following, primal-dual methods by considering the *perturbed* KKT conditions given by

$$F(x, y, \lambda, \sigma, \mu) = \begin{bmatrix} Gx - A^T\lambda + c \\ Ax - y - b \\ \mathcal{Y}\Lambda e - \sigma\mu e \end{bmatrix} = 0, \tag{3.16}$$

where

$$\mathcal{Y} = \mathrm{diag}(y_1, \cdots, y_m),\ \Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_m),\ e = (1, \cdots, 1)^T,$$

and $\sigma \in [0, 1]$. The solutions of (3.16) for all possible values of σ and μ define the *central path*, which is a trajectory that leads to the solution of the quadratic program as $\sigma\mu$ tends to zero.

Interior-Point Methods

By fixing μ and applying Newton's method to (3.16), we obtain the linear system

$$\begin{bmatrix} G & 0 & -A^T \\ A & -I & 0 \\ 0 & \Lambda & \mathcal{Y} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -(Gx - A^T\lambda + c) \\ -(Ax - y - b) \\ -\Lambda \mathcal{Y} e + \sigma \mu e \end{bmatrix}. \tag{3.17}$$

We obtain the next iterate by setting

$$(x^+, y^+, \lambda^+) = (x, y, \lambda) + \alpha(\Delta x, \Delta y, \Delta \lambda),$$

where α is chosen to retain the inequality $(y^+, \lambda^+) > 0$ and possibly to satisfy various other conditions.

Interior-Point Methods

Several enhancements are needed to make the primal-dual iteration effective in practice, such as

- Solving the Primal-Dual System (3.17);
- Step Length Selection;
- Computing corrector steps;

Outline

Perspective and Software

Practical numerical comparisons show that

- The interior-point methods are generally much faster on large problems;
- If a warm start is required, active-set methods are generally preferable;
- Practical projected gradient methods are currently only available for several kinds of QPs whose feasible regions have special structure.

Thanks for your attention!