Quadratic programming

Jitkomut Songsiri

Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University

CUEE

March 6, 2023

Outline

Quadratic function

Quadratic function

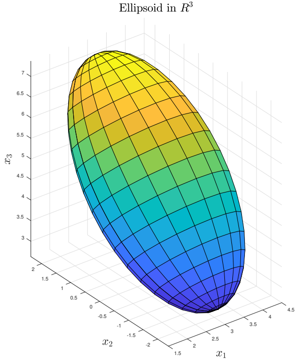
given $P \in \mathbf{R}^{n \times n}, q \in \mathbf{R}^n, r \in \mathbf{R}$, a quadratic function $f : \mathbf{R}^n \to \mathbf{R}$ is of the form

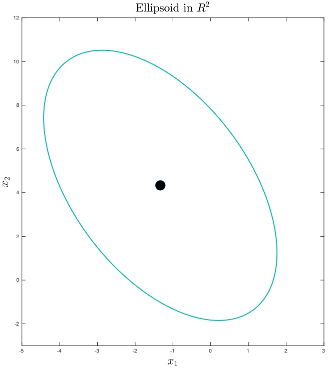
$$f(x) = (1/2)x^T P x + q^T x + r$$

- $x^T P x$ is aka an energy form (due to the quadratic form that appears in the energy/power of some physical variables)
- ✎ verify that $x^T P x = \frac{x^T (P + P^T) x}{2}$; then the energy term only takes the symmetric part of P; hence, we often consider $P \in \mathbf{S}^n$ (P is assumed to be symmetric later on)
- $\nabla f(x) = Px + q$ (derivative of quadratic function becomes linear)
- the contour shape of f depends on the property of P (pdf, indefinite, magnitude of eigenvalues, direction of eigenvectors)

Quadratic function (positive definite)

let $f(x) = (1/2)x^T Px + q^T x$ where $P \succ 0$





since P is invertible, we can complete the square

$$f(x) = (1/2)[(x + P^{-1}q)^T P(x + P^{-1}q) - q^T P^{-1}q]$$

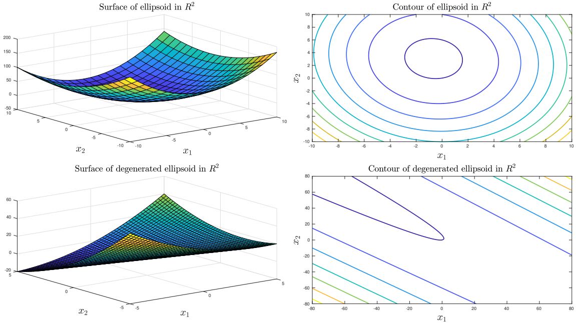
ellipsoid parametrized by P^{-1} with center at $-P^{-1}q$

Quadratic function (positive semidefinite)

let $f(x_1, x_2) = (1/2)(x^T P x) + q^T x$ with $q = (1, -3)$ and two cases of P

$$P = \begin{bmatrix} 2 & 0.1 \\ 0.1 & 1 \end{bmatrix} \succ 0$$

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \succeq 0$$

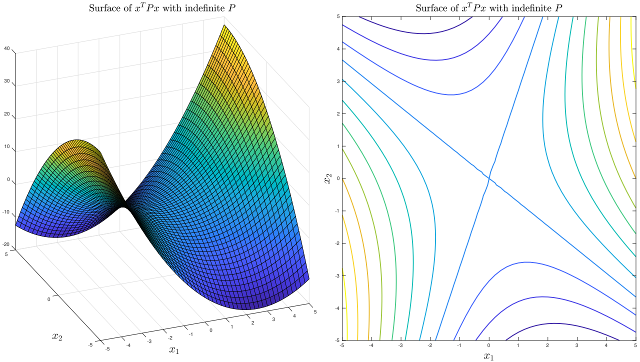


- $P \succ 0$: sublevel set of f is bounded (region inside the ellipsoid)
- $P \succeq 0$: sublevel set of f is unbounded

  (if $x = t(1, -1) \in \mathcal{N}(P)$ then $f(x) = tq^T(1, -1) = 4t \to -\infty$ by choosing $t \to -\infty$)

Quadratic function (indefinite)

let $f(x_1, x_2) = (1/2)(x^T P x) + q^T x$ with $P = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$ (and invertible)



from $f(x) = (1/2)(x + P^{-1}q)^T P(x + P^{-1}q) +$ constant, we can pick t, x such that $x + P^{-1}q = tv, Pv = \lambda^- v, t \to \infty$; hence, $f(x) = t^2 \lambda^- \|v\|^2 \to -\infty$

f can be unbounded below along some direction of x

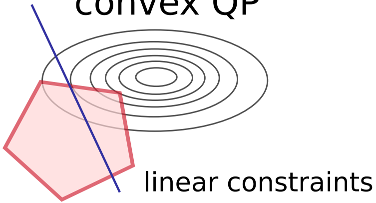
Formulation

Standard form

a **quadratic program (QP)** is in the form

$$\begin{array}{ll} \text{minimize} & (1/2)x^TPx + q^Tx \\ \text{subject to} & Gx \preceq h \\ & Ax = b, \end{array}$$

where $P \in \mathbf{S}^n, G \in \mathbf{R}^{m \times n}$ and $A \in \mathbf{R}^{p \times n}$



example: constrained least-squares

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2^2 \\ \text{subject to} & l \preceq x \preceq u \end{array}$$

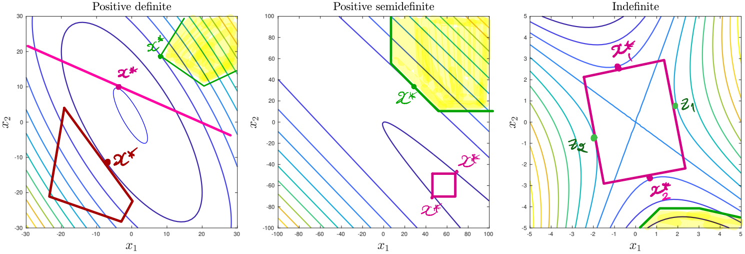
QP has **linear** constraints

Properties of QP

- an unconstrained QP is unbounded below if P is not positive definite
- an unconstrained QP has a unique solution: $x = -P^{-1}q$ when $P \succ 0$
- a QP is a convex problem if P is positive semidifinite definite
  - if $P \succeq 0$ then a local minimizer $x^\star$ is a global minimizer (by convexity)
  - if $P \succ 0$ then $x^\star$ is a *unique* global solution (by strictly convexity)
- the feasible set (polyhedron) may be empty (hence, the problem is infeasible)
- the feasible set can be unbounded (but if $P \succ 0$ it implies boundedness)
- solution of a QP may not be at a vertex
- the dual of a QP is also a QP

Contour of quadratic objective

consider three cases of P and different feasible sets



verify the location of the optimal solution for each constraint set

- left: a bounded set, a line, an unbounded feasible set
- middle: bounded and unbouded feasible sets, while f is unbounded below
- right: a bounded feasible set, while f is unbounded below and above

Equality-constrained QP

assume a full row rank matrix $A \in \mathbf{R}^{p\times n}$ and $P \succ 0$ on the **nullspace** of A

$$\underset{x}{\text{minimize}} \;\; (1/2)x^T P x - q^T x \quad \text{subject to} \;\; Ax = b$$

- it can be shown that $K = \begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}$ is non-singular (called KKT matrix)
- the zero-gradient of Lagrangian condition is the system of $n+p$ equations

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} q \\ b \end{bmatrix}$$

  has a unique solution $(x^\star, \lambda^\star)$
- $x^\star$ is the unique **global** solution proof in Thm 16.2, Nocedral book

Proof

suppose the KKT matrix is singular, $\exists z = (x, v) \neq 0$ such that $Kz = 0$, hence

- $Ax = 0$ (x lies in the nullspace of A) and $Px + A^T v = 0$
- $z^T K z = 0$ and this gives

$$z^T \begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} z = x^T P x + 2v^T A x = x^T P x = 0$$

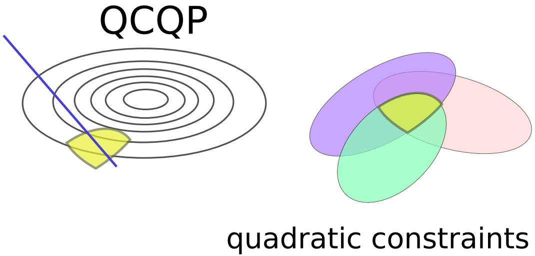
- but $P \succ 0$ for all $y \in \mathcal{N}(A)$, hence $x^T P x = 0$ only holds when $x = 0$
- when $x = 0$, we conclude from $Px + A^T v = 0$ that $A^T v = 0$
- but A is full row rank (making $A^T v$ full column rank), we conclude that $v = 0$
- this leads to a contradiction, $(x, v) = 0$ so K can't be singular

QCQP

a **quadratically constrained quadratic program (QCQP)** is in the form

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P_0 x + q_0^T x \\ \text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m \\ & Ax = b, \end{array}$$

assume P_i's are positive semidefinite, $G \in \mathbf{R}^{m \times n}$ and $A \in \mathbf{R}^{p \times n}$



QCQP has both **linear and quadratic** constraints

Minimizing linear objective under a quadratic constraint

a special case of QCQP where the objective is linear

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & (x-d)^T P^{-1}(x-d) \leq 1 \end{array}$$

where $P \succ 0, d \in \mathbf{R}^n$ are given parameters

- make change of variable: $z = P^{-1/2}(x-d)$

$$\text{minimize} \quad \tilde{c}^T z + g \quad \text{subject to} \quad z^T z \leq 1$$

  where $\tilde{c} = P^{1/2}c$ and $g = c^T d$ is a constant term

- the equivalent problem has a closed-form solution:

$$z^\star = -\frac{\tilde{c}}{\|\tilde{c}\|_2} = -\frac{P^{1/2}c}{\|P^{1/2}c\|_2} \quad \Longrightarrow \quad x^\star = P^{1/2}z^\star + d = -\frac{Pc}{\sqrt{c^T P c}} + d$$

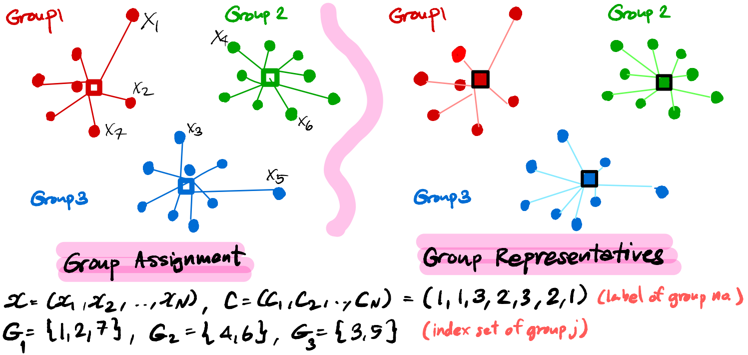
Applications

Applications of quadratic programming

- unconstrained QP
  - least-squares
  - optimizing group representative step in k-mean clustering
- support vector machine
- control systems
- inverse problem (medical imaging, signal processing)
- least-squares with constraints (lasso and others)
- portfolio optimization

k-mean clustering

define c_i the group number of x_i (data) and a group assignment $G_j = \{i \mid c_i = j\}$



after the k groups are assigned, optimizing the group representative (z_j) is to minimize

$$J^{\text{clust}} = J_1 + \cdots + J_k, \quad J_j = (1/N) \sum_{i \in G_j} \|x_i - z_j\|_2^2$$

- updating group representatives is an unconstrained QP in $z = (z_1, \dots, z_k)$
- the solution z_j is the mean (or centroid) of x_i in jth group

$$z_j = \frac{1}{|G_j|} \sum_{i \in G_j} x_i$$

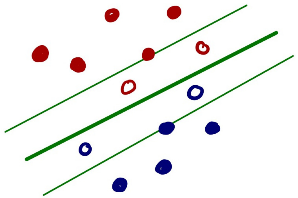
- the scheme of k-mean algorithm consists of
  - partition the data x into k groups (not optimization problem)
  - update the representatives: unconstrained QP (closed-form solution)

Soft-margin SVM

problem parameters: $x_i \in \mathbf{R}^n$ and $y_i \in \mathbf{R}$ for $i = 1, \ldots, N, \lambda > 0$

optimization variables: $w \in \mathbf{R}^n, b \in \mathbf{R}, z \in \mathbf{R}^N$

$$\begin{array}{ll} \text{minimize} & (1/2)\|w\|_2^2 + \lambda \mathbf{1}^T z \\ \text{subject to} & y_i(x_i^T w + b) \geq 1 - z_i, \quad i = 1, 2, \ldots, N \\ & z \succeq 0 \end{array}$$



- data are classified by separating hyperplane with maximized margin
- z_i is called a **slack variable**, allowing some of the hard constraints to be relaxed
- the problem has (convex) quadratic objective and linear constraints (QP)

Tracking problem

design problem: find $u(t)$ for $t = 1, 2, \dots, T$ to drive the linear system

$$x(t+1) = Ax(t) + Bu(t), \quad y(t) = Cx(t), \quad x(0) = 0$$

so that $y \approx y_{\mathrm{ref}}$

the relationship between y and u is

$$y(t) = CA^{t-1}Bu(0) + CA^{t-2}Bu(1) + \cdots + CABu(t-2) + CBu(t-1) + Du(t)$$

and can be arranged into vector form as

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(T) \end{bmatrix} = \begin{bmatrix} CB & 0 & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{T-1}B & \cdots & CAB & CB \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(T-1) \end{bmatrix} \triangleq \quad y_T = Hu_T \qquad (1)$$

Specifications in control tracking

four types of constraints based on specification of u can be cast as a QP

let the optimization variable be $u^T = (u(1), \ldots, u(T))$

- trade-off between tracking and energy of u

$$\underset{u_T}{\text{minimize}} \quad \|Hu_T - y_{\text{ref}}\|_2^2 + \gamma \|u_T\|_2^2 \tag{2}$$

  (unconstrained, closed-form solution, depends on the property of H)

- magnitude of u must be bounded, $|u| \leq u_{\max}$

$$\underset{u_T}{\text{minimize}} \quad \|Hu_T - y_{\text{ref}}\|_2^2 \quad \text{subject to} \quad -u_{\max} \preceq u_T \preceq u_{\max} \tag{3}$$

Specifications in control tracking

- the control signal does not change too rapidly, $|u(k) - u(k-1)|$ is small

$$\begin{array}{ll} \text{minimize}_{u_T} & \|Hu_T - y_{\mathrm{ref}}\|_2^2 + \gamma\|Du_T\|_2^2 \\ \text{subject to} & -u_{\max} \preceq u_T \preceq u_{\max} \end{array} \quad (4)$$

where $D : \mathbf{R}^T \to \mathbf{R}^{T-1}$ is the difference matrix

- rate of change in u is bounded

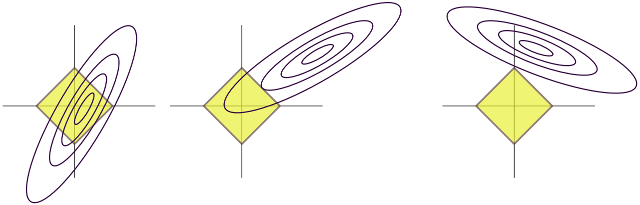
$$\begin{array}{ll} \text{minimize}_{u_T} & \|Hu_T - y_{\mathrm{ref}}\|_2^2 \\ \text{subject to} & -u_{\max} \preceq u_T \preceq u_{\max} \\ & -d_{\max}\mathbf{1} \preceq Du_T \preceq d_{\max}\mathbf{1} \end{array} \quad (5)$$

Lasso as a convex QP

a lasso or basis pursuit is the problem

$$\underset{x}{\text{minimize}}\ \|Ax - b\|_2^2 \quad \text{subject to} \quad \|x\|_1 \leq t$$

minimizing the residual norm while keeping norm of x small (controlled by t)



this can be cast as a convex QP (since $A^T A \succeq 0$) with variables $x, u \in \mathbf{R}^n$

$$\begin{array}{ll} \text{minimize} & x^T A^T A x - 2b^T A x \\ \text{subject to} & -u \preceq x \preceq u \\ & \mathbf{1}^T u \leq t \end{array}$$

ℓ_1-regularized least-squares

an ℓ_1-regularized least-squares (Lagrangian form of lasso)

$$\underset{x}{\text{minimize}} \; \|Ax - b\|_2^2 + \gamma \|x\|_1$$

QCQP formulation:

using the epigraph form, we can formulate the problem as

$$\begin{array}{ll} \text{minimize} & t + \gamma \mathbf{1}^T u \\ \text{subject to} & x^T A^T x - 2b^T Ax + b^T b \leq t \\ & -u \preceq x \preceq u \end{array}$$

with variables $x, u \in \mathbf{R}^n$ and $t \in \mathbf{R}$

QP formulation: note that we can write x as

$$x = u - v, \quad u, v \succeq 0 \quad \Rightarrow \quad |x| = u + v \quad \text{(all elementwise)}$$

u and v are positive and negative parts of x, respectively

$$\|x\|_1 = \sum_k |x_k| = \mathbf{1}^T(u + v)$$

the problem can be formulated as a QP

$$\begin{array}{ll} \text{minimize} & \|Ax - y\|_2^2 + \gamma \mathbf{1}^T(u + v) \\ \text{subject to} & x = u - v \\ & u \succeq 0, \quad v \succeq 0 \end{array}$$

with variables $x, u, v \in \mathbf{R}^n$

Markowitz portfolio optimization

setting:

- $r = (r_1, r_2, \ldots, r_n) \in \mathbf{R}^n$; r_i is the (random) return of asset i
- the return has the mean $\bar{r}$ and covariance Σ

optimization variable: $x \in \mathbf{R}^n$ where x_i is the portion to invest in asset i

problem parameters: $\Sigma \succeq 0, \bar{r} \in \mathbf{R}^n, \gamma > 0$

$$\begin{array}{ll} \text{minimize} & -\bar{r}^T x + \gamma x^T \Sigma x \\ \text{subject to} & x \succeq 0, \quad \mathbf{1}^T x = 1 \end{array}$$

- $\mathbf{var}(r^T x) = x^T \Sigma x$ is the risk of the portfolio
- the goal is to maximize the expected return while minimize the risk
- γ is the risk-aversion parameter controlling the trade-off

Risk minimization with fixed return

setting: consider returns of n assets in T periods

- $R \in \mathbf{R}^{T \times n}$: R_{ij} is the gain of asset j in period i (%)
- $w \in \mathbf{R}^n$: asset allocation (or weight) where $\mathbf{1}^T w = 1$
- $r \in \mathbf{R}^T$: r_i is the return (of all assets) in period i, so $r = Rw$
- total portfolio value in period t is

$$V_t = V_1(1 + r_1)(1 + r_2) \cdots (1 + r_{t-1})$$

  and can be approximated when r_t is small as $V_{T+1} \approx V_1 + T\,\mathbf{avg}(r) V_1$
- unlike Markowitz that used statistical property of the returns, here we use a set of actual (or realized) returns
- as seen in Markowitz formulation, w that minimize risk for a given return is called **Pareto optimal**

Risk minimization with fixed return

goal: fix the return to a value ρ and minimize the risk over all portfolios

- the portfolio return is given by $\mathbf{avg}(r) = (1/T)\mathbf{1}^T(Rw) \triangleq \mu^T w = \rho$
- the risk is $\mathbf{var}[r] = (1/T)\|r - \mathbf{avg}(r)\|^2 = (1/T)\|r - \rho\mathbf{1}\|^2$

the problem of minimizing the risk with return ρ is

$$\begin{array}{ll} \text{minimize} & \|Rw - \rho\mathbf{1}\|^2 \\ \text{subject to} & \begin{bmatrix} \mathbf{1}^T \\ \mu^T \end{bmatrix} w = \begin{bmatrix} 1 \\ \rho \end{bmatrix} \end{array}$$

with variable $w \in \mathbf{R}^n$ and parameters R, ρ, μ

(no non-negative constraint in w – this gives quadratic programming with linear equality)

Algorithms

Available methods

- active set method for convex QPs
- interior-point methods
- conjugate gradient (solving the reduced problem of equality-constrained QP)
- ellipsoid method (for convex programs): generate a sequence of ellipsoids that are guaranteed to contain the minizer
- gradient projection (for QP if the polyhedron is simple)
- many solvers and packages in the market

MATLAB: `quadprog` use trust-region-reflective or interior-point

Python (convex QP and QCQP): `cvxopt`

Active-set methods for convex QP

- standard form
- algorithm outline
- update working set (that contains active constraints)
- optimality condition

QP standard form for active-set methods

we consider the standard form of convex QP with inequality constraints:

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x \\ \text{subject to} & a_i^T x = b_i, \quad i \in \mathcal{E} \\ & a_i^T x \geq b_i, \quad i \in \mathcal{I} \end{array}$$

- the **active set** $\mathcal{A}(x)$ consists of i of the constraints for which equality holds at x

$$\mathcal{A}(x) = \{ \, i \in \mathcal{E} \cup \mathcal{I} \mid a_i^T x = b_i \, \}$$

  (we typically don't have knowledge of $\mathcal{A}(x^\star)$)
- at iteration k when updating x_k, define $\mathcal{W}_k$ as the **working set** which contains $i \in \mathcal{E}$ and *some* indices from $\mathcal{I}$ that inequalities are imposed as equalities
- it is required that a_i's for $i \in \mathcal{W}_k$ are linearly independent

Algorithm outline

the updates rely on subproblems that solve QP with linear equalities

1. given an initial **feasible** point x_0
2. the update takes the form of $x_{k+1} = x_k + \alpha_k s_k$
3. at iterate x_k, we can determine $\mathcal{W}_k$
4. finding s_k is to solve QP subproblem with **equality constraints** for $i \in \mathcal{W}_k$ (this is an easy problem – refer to page 12)
5. update $\mathcal{W}_k$ by either add or remove i corresponding to inequality constraints
6. the update terminates when $s_k = 0$ and KKT conditions are satisfied

QP subproblem to find the search direction

given x_k and the working set $\mathcal{W}_k$, we solve the QP

$$\text{minimize} \quad (1/2)s^T P s + (Px_k + q)^T s \quad \text{subject to} \quad a_i^T s = 0, \quad i \in \mathcal{W}_k$$

and the optimal solution s is then assigned to search direction s_k

- the constraints corresponding to $\mathcal{W}_k$ are regarded as equalities where all other constraints are temporarily disregarded
- we solve QP subproblem using the technique on page 12 (solve KKT system)
- using $L(s, \lambda) = (1/2)s^T P s + (Px_k + q)^T s - \sum_i \lambda_i a_i^T s$, the KKT system is

$$\begin{bmatrix} P & -A_w^T \\ A_w & 0 \end{bmatrix} \begin{bmatrix} s \\ \lambda \end{bmatrix} = \begin{bmatrix} -(Px_k + q) \\ 0 \end{bmatrix} \tag{6}$$

(A_w contains rows of a_i^T for $i \in \mathcal{W}_k$)

Determining stepsize

to update $x_{k+1} = x_k + \alpha_k s_k$, we check the feasibility of x_{k+1}

- if $\alpha_k = 1$ makes x_{k+1} feasible (to all constraints) then set $x_{k+1} = x_k + s_k$; otherwise, find an appropriate value of $\alpha \in [0, 1]$
- as we only need to check feasibility of constraints for $i \notin \mathcal{W}_k$
  - if $a_i^T s_k \geq 0$ then we can use any $\alpha_k \geq 0$ because x_{k+1} is always feasible

$$a_i^T(x_k + \alpha_k s_k) = a_i^T x_k + \alpha_k a_i^T s_k \geq a_i^T x_k \geq b_i$$

  - if $a_i^T s_k < 0$ for some $i \notin \mathcal{W}_k$, we make $a_i^T(x_k + \alpha_k s_k) \geq b_i$ only if we choose

$$\alpha_k \leq \frac{b_i - a_i^T x_k}{a_i^T s_k}$$

(there can be many i's that $a_i^T s_k < 0$, so we pick smallest α_k in $[0, 1]$)

Blocking constraints

in conclusion, when $a_i^T s_k < 0$ for some $i \notin \mathcal{W}_k$, we set

$$\alpha_k = \min\left(1, \min_{i \notin \mathcal{W}_k, a_i^T s_k < 0} \frac{b_i - a_i^T x_k}{a_i^T s_k}\right)$$

- **blocking constraints** are the constraints i for which the minimum occurs
  - if $\alpha_k < 1$, step along s_k was blocked by some $i \notin \mathcal{W}_k$, so we adjust by $\mathcal{W}_{k+1} := \mathcal{W}_k \cup$ blocking constraints
  - if $\alpha_k = 1$, then no blocking constraints and $\mathcal{W}_{k+1} = \mathcal{W}_k$
- iterate k until we find that $s_k \triangleq \hat{s} = 0$ (with the current working set $\hat{\mathcal{W}}$)
- the KKT condition of QP subproblem on page 35 suggests that

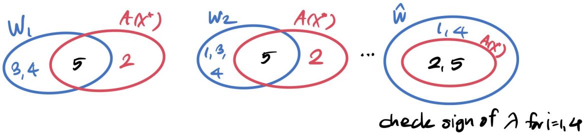
$$P\hat{x} + q = \sum_{i \in \hat{\mathcal{W}}} a_i \hat{\lambda}_i$$

Checking optimality

KKT conditions of the original QP problem on page 33

primal feasibility: $a_i^T x^\star = b_i, \forall i \in \mathcal{A}(x^\star), \quad a_i^T x^\star \geq b_i, \forall i \in \mathcal{I} \backslash \mathcal{A}(x^\star)$

zero-gradient: $Px^\star + q - \sum_{i \in \mathcal{A}(x^\star)} \lambda_i^\star a_i = 0,$ **dual feasibility:** $\lambda_i^\star \geq 0, \forall i \in \mathcal{I} \cap \mathcal{A}(x^\star)$



conditions obtained from $\hat{x}, \hat{\lambda}$

- $P\hat{x} + q - \sum_{i \in \hat{\mathcal{W}}} \hat{\lambda}_i a_i - \sum_{i \notin \hat{\mathcal{W}}} 0 \cdot a_i = 0$
- $a_i^T \hat{x} = b_i, \forall i \in \mathcal{A}(\hat{x})$
- $a_i^T \hat{x} \geq b_i, \quad \forall i \in \mathcal{I} \backslash \mathcal{A}(\hat{x})$ because $a_i^T \hat{x} = b_i$ for $i \notin \mathcal{A}(\hat{x})$ but $i \in \hat{\mathcal{W}}$
- it's left to check if $\hat{\lambda}$ for all $i \in \mathcal{I} \cap \hat{\mathcal{W}}$

Sign of Lagrange multipliers

we examine the sign of $\hat{\lambda}_i$ for $i \in \mathcal{I} \cap \hat{\mathcal{W}}$

- if $\hat{\lambda}_i \succeq 0$ then $\hat{\lambda}$ is dual feasible and $\hat{x}$ is optimal (satisfying all KKT conditions)
- if $\hat{\lambda}_j < 0$ for some $j \in \mathcal{I} \cap \hat{\mathcal{W}}$
  - find j that $\hat{\lambda}_j$ is most negative
  - remove j from the working set: $\mathcal{W}_{k+1} := \mathcal{W}_k \backslash j$

  (the decreasing rate of objective function when one constraint is removed is proportional to Lagranger multiplier of that constraint)

then continue iteration k and solve the QP subproblem

Algorithm: active-set method for convex QP

Require: tolerance = 1e-5 , maxiter = 50

Initialize: feasible point x_0
for $k = 1$: maxiter **do**
  solve QP subproblem on page 35 to find s_k
  if $\|s_k\| \leq$ tolerance **then**
    compute $\hat{\lambda}$ with $\hat{\mathcal{W}} = \mathcal{W}_k$
    if $\hat{\lambda}_i \geq 0$ for all $i \in \mathcal{W}_k \cap \mathcal{I}$ **then**
      stop with solution $x^\star = x_k$
    else
      $j = \mathrm{argmin}_{j \in \mathcal{W}_k \cap \mathcal{I}} \hat{\lambda}_j$
      $x_{k+1} := x_k$; $\mathcal{W}_{k+1} := \mathcal{W}_k \backslash \{j\}$
    end if
  else
    compute α_k from page 37
    $x_{k+1} := x_k + \alpha_k s_k$
    if there are blocking constraints **then**
      obtain $\mathcal{W}_{k+1}$ by adding one of blocking constraints to $\mathcal{W}_k$
    else
      $\mathcal{W}_{k+1} := \mathcal{W}_k$
    end if
  end if
end for
return x_k

References

1. L. Vandenberghe, *Lecture notes on Optimization Methods for Large-Scale Systems*, EE236C, UCLA
2. D.P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1999
3. G. Calafiore and L. El Ghaoui, *Optimization Models*, Cambridge University Press, 2014
4. S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004
5. J. Nocedal and S.J. Wright, *Numerical Optimization*, 2nd edition, Springer, 2006
6. S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least squares*, Cambridge University Press, 2018
7. J. Songsiri, *System Identification*, Chula Press, 2022