

# Predictive Control for Linear and Hybrid Systems

Borrelli

Bemporad

Morari

Alexander Brown

## Main Concepts

### Optimization Problems

An optimization problem is generally formulated as

$$\inf_{z \in S \subseteq Z} f(z)$$

Where  $\inf(\cdot)$  resembles finding the optimal value within the subset.

Solving this problem means to compute the least possible cost  $f^*$ .

$$f^* = \inf_{z \in S} f(z)$$

The number  $f^*$  is the **optimal value** of  $\inf_{z \in S \subseteq Z} f(z)$ , i.e.:

$$f(z) \geq f(z^*) = f^* \quad \forall z \in S, \text{ with } z^* \in S$$

### Continuous Problems

#### *Nonlinear mathematical program*

$$\begin{array}{ll} \inf_z & f(z) \\ \text{subj. to} & g_i(z) \leq 0 \quad \text{for } i = 1, \dots, m \\ & h_j(z) = 0 \quad \text{for } j = 1, \dots, p \\ & z \in Z \end{array}$$

A point  $\bar{z} \in \mathbb{R}^s$  is **feasible** for the continuous optimization problems if:

1. it belongs to  $Z$
2. It satisfies the inequality and equality constraints

$\bar{z}$  is also a *local optimizer* or *local minimizer* if it satisfies the constraints above as well as  $\|z - \bar{z}\| \leq R$  where  $R$  defines a region (think of stability definitions Linear Multi-Variable Control).

**Active, Inactive, and Redundant Cases:** Consider a feasible point.  $g_i(\bar{z}) = 0$  is active,  $g_i(\bar{z}) < 0$  is inactive, and a constraint is redundant if removing the constraint does not change the feasible set  $S$ .

**Eliminating Equality Constraints:** The simplest way to remove equality constraints is to replace them with two inequalities:

$$\begin{array}{l} h_i(z) = 0 \\ \therefore \text{is replaced by} \\ h_i(z) \leq 0 \\ -h_i(z) \leq 0 \end{array}$$

### Integer and Mixed-Integer Problems

If the optimization problem

$$\inf_{z \in S \subseteq Z} f(z)$$

is finite, then the optimization problem is called *combinatorial* or *finite*. If  $Z \subseteq 0, 1^s$ , then the problem is said to be *integer*. If  $Z$  is a subset of the Cartesian product of an integer set and real Euclidian space, then the problem is said to be *mixed-integer*. The standard form of a mixed-integer nonlinear program is:

$$\begin{array}{lll}
\text{inf}_{[z_c, z_b]} & f(z_c, z_b) & \\
\text{subj. to} & g_i(z_c, z_b) \leq 0 & \text{for } i = 1, \dots, m \\
& h_j(z_c, z_b) = 0 & \text{for } j = 1, \dots, p \\
& z_c \in \mathbb{R}^{s_c}, z_b \in [0, 1]^{s_b} &
\end{array}$$

## Convexity

The standard optimization problem is said to be convex if the cost function  $f$  is convex on  $Z$  and  $S$  is a convex set. A fundamental property of convex optimization problems is that local optimizers are also global optimizers (think calculus optimization: min/max of parabola is global).

### Theorem 1.1

Consider a convex optimization problem and let  $\bar{z}$  be a local optimizer. Then  $\bar{z}$  is a global optimizer.

## Optimality Conditions

### Optimality Conditions For Unconstrained Problems

#### Theorem 1.2 (Necessary Condition):

Suppose that  $f : \mathbb{R}^s \rightarrow \mathbb{R}$  is differentiable at  $\bar{z}$ . If there exists a vector  $d$  such that  $\nabla f(\bar{z})'d < 0$ , then there exists a  $\delta > 0$  such that  $f(\bar{z} + \lambda d) < f(\bar{z})$  for a  $\lambda \in (0, \delta)$ .

#### Theorem 1.3 (Sufficient Condition):

Suppose that  $f : \mathbb{R}^s \rightarrow \mathbb{R}$  is twice differentiable at  $\bar{z}$ . If  $\nabla f(\bar{z}) = 0$  and the Hessian  $(\nabla^2 f(\bar{z}))$  of  $f(z)$  at  $\bar{z}$  is positive definite, then  $\bar{z}$  is a local minimizer.

#### Theorem 1.4: (Necessary and Sufficient Condition):

Supposed that  $f : \mathbb{R}^s \rightarrow \mathbb{R}$  is differentiable at  $\bar{z}$ . If  $f$  is convex, then  $\bar{z}$  is a global minimizer iff  $\nabla f(\bar{z}) = 0$

## Lagrange Duality Theory

Consider the optimality problem. Any feasible point  $\bar{z}$  provides an upper bound to the optimal value  $f(\bar{z}) \geq f^*$  ( $f^*$  being the optimal value). The Lagrange Duality Theory generates a lower boundary for  $f^*$ .

Starting from the same problem, we construct another problem with different variables and constrains. In other words, from the primal problem, we will develop the dual problem.

$$L(z, u, v) = f(z) + u_1 g_1(z) + \dots + u_m g_m(z) + v_1 h_1(z) + \dots + v_p h_p(z)$$

The primal problem is written as the problem:

$$\sup_{(u,v), u \geq 0} \inf_{z \in Z} L(z, u, v) \leq \inf_{s \in Z, g(z) \leq 0, h(z) = 0} f(z)$$

Suppose we define  $d(u, v) = \inf_{z \in Z} L(z, u, v)$ , then we can restate the previous problem as:

$$\sup_{(u,v), u \geq 0} d(u, v)$$

where its optimal solution, assuming it exists, is denoted by  $(u^*, v^*)$ . The dual cost  $d(u, v)$  is the optimal solution of the *unconstrained optimization problem*. A point  $(u, v)$  is dual feasible if  $u \geq 0$  and  $d(u, v) \geq 0$ .  $d(u, v)$  is *always a concave function* since it is the pointwise infimum of a family of affine functions of  $(u, v)$ . This implies that the *dual problem is a convex optimization problem even if the original problem is not convex*. In general, however, the solution to the dual problem is only a lower bound of the primal problem (known as *weak duality*). In other words there is a gap between the primal and dual solutions.

The difference  $f^* - d^*$  is the duality gap, where  $f^*$  and  $d^*$  are the optimal solutions of both the primal and dual problem respectively.

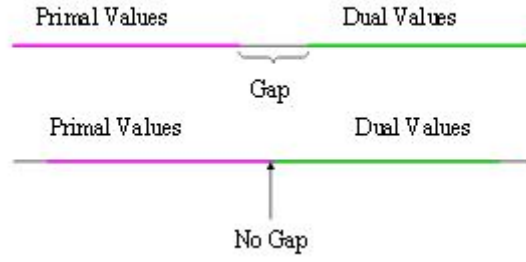


Figure 1: Weak and Strong Duality

## Strong Duality and Constraint Qualifications

If the cost of the dual problem is equal to the cost of the primary problem ( $d^* = f^*$ ) then we have **strong duality**. In general, strong duality does not hold, even from convex primal problems (but can only hold if some suitable convexity conditions are satisfied). *Constraint Qualifications are conditions on the constraint functions which imply strong duality for convex problem.*

### Slater's Condition:

Consider the optimality problem stated above. There exists a  $\hat{z} \in \mathbb{R}^s$  which belongs to the relative interior of the problem domain  $Z$ , which is feasible ( $g(\hat{z}) \leq 0, h(\hat{z}) = 0$ ) and for which  $g_j(\hat{z}) < 0$  for all  $j$  for which  $g_j$  is not an affine function.

### Theorem 1.5 (Slater's Theorem):

Consider the primal and its dual problem. If the primal problem is convex, Slater's condition holds and  $f^*$  is bounded then  $d^* = f^*$ .

Recall that this is for the unconstrained problem.

## Certificate of Optimality

Without knowing the exact value of  $f^*$  we can give a bound on how suboptimal a given feasible point is. In fact, if  $z$  is primal feasible and  $(u, v)$  is dual feasible then  $d(u, v) \leq f^* \leq f(z)$ . Therefore  $z$  is  $\epsilon$ -suboptimal, with  $\epsilon = f(z) - d(u, v)$  (primal-dual gap). Therefore many algorithms utilize

$$f(z_k) - d(u_k, v_k) < \epsilon_k$$

for the termination condition. This condition at the worst supplies  $\epsilon$ -suboptimal. If strong duality holds, the condition can be met for arbitrarily small tolerances (will converge to the exact solution).

## Complimentary Slackness

The complimentary slackness is represented as:

$$u_i * g_i(z^*) = 0, \quad i = 1, \dots, m$$

and it can be interpreted as follows. If the  $i$ -th inequality constraint of the primal problem is inactive at the optimum ( $g_i(z^*) < 0$ ), then the  $i$ -th dual optimizer has to be zero ( $u_i = 0$ ). Vice versa, if the  $i$ -th dual optimizer is different from zero ( $u_i > 0$ ), then the  $i$ -th constraint is active at the optimum ( $g_i(z^*) = 0$ ).

## Karush-Kuhn-Tucker Conditions

The Primal and dual optimal pair  $z^*, (u^*, v^*)$  of an optimization problem with differentiable cost and constraints and zero duality gap, have to satisfy the following conditions:

$$\begin{aligned} \nabla f(z^*) + \sum_i u_i * \nabla g_i(z^*) + \sum_j v_j * \nabla h_j(z^*) &= 0 \\ u_i * g_i(z^*) &= 0 \quad i = 1, \dots, m \\ u_i &\geq 0 \quad i = 1, \dots, m \\ g_i(z^*) &\leq 0 \quad i = 1, \dots, m \\ h_j(z^*) &= 0 \quad j = 1, \dots, p \end{aligned}$$

Where:

- $g_i(z^*) \leq 0$  and  $h_j(z^*) = 0$  are the primal feasibility conditions
- $u_i^* \geq 0$  is the dual feasibility condition
- $u_i^* g_i(z^*) = 0$  is the complimentary slack condition

**Theorem 1.6**

Consider the primal optimization problem. Also suppose the problem is convex and that cost and constraints  $f$ ,  $g_i$ , and  $h_i$  are differentiable and feasible  $z^*$ . If the problem satisfies Slater's condition then  $z^*$  is optimal iff there are  $(u^*, v^*)$  that, together with  $z^*$ , satisfy the KKT conditions.

**If a convex optimization problem with differentiable objective and constraint functions has linearly independent set of active constraint gradients, then the KKT conditions provide necessary and sufficient conditions for optimality.**

**Theorem 1.7**

Consider the primal problem and let  $Z$  be a nonempty open set of  $\mathbb{R}^s$ . Let  $z^*$  be a feasible solution and  $A = \{i : g_i(z^*) = 0\}$  be the set of active constraints at  $z^*$ . Suppose cost and constraints  $f$  and  $g_i$  are differentiable at  $z^*$  for all  $i$  and that  $h_j$  are continuously differentiable at  $z^*$  for all  $j$ . Further, suppose that  $\nabla g_i(z^*)$  for  $i \in A$  and  $\nabla h_j(z^*)$  for  $j = 1, \dots, p$ , are linearly independent. If  $z^*$ ,  $(u^*, v^*)$  are primal and dual optimal points, then they satisfy the KKT conditions. In addition, if the primal problem is convex, the  $z^*$  is optimal iff there are  $(u^*, v^*)$  that, together with  $z^*$ , satisfy the KKT conditions.

## Linear and Quadratic Optimization

### Linear Programming

When the cost and the constraints of the continuous optimization problem are affine, then the problem is called a *linear program*. The most general form of a LP is:

$$\begin{array}{ll} \inf_z & c'z \\ \text{subject to} & Gz \leq w \end{array}$$

Two other common forms are:

$$\begin{array}{ll} \inf_z & c'z \\ \text{subject to} & Gz \leq w \\ & Az = b \end{array}$$

and

$$\begin{array}{ll} \inf_z & c'z \\ \text{subject to} & Gz \leq w \\ & z \geq 0 \end{array}$$

By standard manipulations, it is always possible to convert one of the three forms into the other.

### Geometric Interpretation and Solutions Properties

Three cases can occur when solving an LP

1. The solution is unbounded  $f^* = -\infty$
2. The solution is bounded  $f^* > -\infty$  and the optimizer is unique ( $z^* = Z^*$  is a singleton)
3. The solution is bounded and there are multiple optima  $Z^*$  is a subset of  $\mathbb{R}^s$  which can be either bounded or unbounded.

### Dual of LP

Consider the LP

$$\begin{array}{ll} \inf_z & c'z \\ \text{subject to} & Gz \leq w \end{array}$$

The dual problem is written as:

$$\begin{array}{ll} \sup_u & -u'w \\ \text{subject to} & -G'u = c \\ & u \geq 0 \end{array}$$

## KKT condition for LP

$$\begin{array}{l} G'u + c = 0 \\ u_i(G_i z - w_i) = 0 \quad i = 1, \dots, m \\ u \geq 0 \\ Gz - w \leq 0 \end{array}$$

## Active Constraints and Degeneracies

Suppose we have a matrix of constraints  $G$ , we can decompose that matrix into active and inactive sections (think observable decomposition from Multi-Variable Control):

$$\begin{array}{l} A(z) = i \in I : G_i z = w_i \\ NA(z) = i \in I : G_i z < w_i \end{array}$$

**Definition 2.1:** (Linear Independence Constraint Qualification (LICQ))

We say the LICQ holds at  $z^*$  if the matrix  $G_{A(z^*)}$  has full row rank.

**Lemma 2.1**

Assume that the feasible set  $P$  of a primal LP is bounded. If the LICQ is violated at  $z_1^* \in Z^*$  then there exists a  $z_2^* \in Z^*$  such that  $|A(z_2^*)| > s$ .

Therefore, violation of LICQ is equivalent to having more than  $s$  constraints active at an optimal vertex.

**Definition 2.2**

The primal LP is said to be primal degenerate if there exists a  $z^* \in Z^*$  such that LICQ does not hold at  $z^*$

**Definition 2.3**

The primal LP is said to be dual degenerate if the dual problem is primal degenerate

**Lemma 2.2**

If the primal problem is not degenerate then the dual problem has a unique optimizer. If the dual problem is not degenerate, then the primal problem has a unique optimizer.

## Quadratic Programming

The continuous optimization problem is called a quadratic program (QP) if the constraint functions are affine and the cost function is a convex quadratic function. QPs can be represented in the general form:

$$\begin{array}{ll} \min_z & 1/2 z' H z + q' z + r \\ \text{subject to} & Gz \leq w \end{array}$$

or

$$\begin{array}{ll} \min_z & 1/2 z' H z + q' z + r \\ \text{subject to} & Gz \leq w \\ & Az = b \end{array}$$

## Geometric Interpretation and Solution Properties

The two dimensional geometric interpretation is depicted as ellipsoid level curves (calculus 3: partial derivatives being visualized). All the points  $z$  belonging to both the ellipsoid (cost function) and the polyhedron (feasible set) are feasible points with an associated cost  $k$ . The smaller the ellipsoid, the smaller is its cost  $k_i$ . This amounts to finding a feasible  $z$  which belongs to the level curve with the smallest  $k_i$ . Since  $H$  is strictly positive definite, the QP cannot have multiple optima nor unbounded solutions. If  $P$  (the feasible set) is not empty, the optimizer  $z$  is unique. This results in two cases:

1. The optimizer lies strictly inside the feasible polyhedron
2. The optimizer lies on the boundary of the feasible polyhedron

## Dual of QP

Because of the required convexity of this sort of problem, it is both necessary and sufficient to state that the optimality is found when the gradient is zero.

$$\begin{aligned} Hz + q + G'u &= 0 \\ \therefore \\ z &= -H^{-1}q'z + u'(Gz - w) \end{aligned}$$

Substituting the solution above into the dual cost (i.e. the Lagrangian of the Quadratic Program  $d(u) = \min_z 1/2z'H z + z'z + u'(Gz - w)$ ):

$$\begin{aligned} \min_u \quad & d(u) = -u'(GH^{-1}G')u - u'(w + GH^{-1}q) - 1/2q'H^{-1}q \\ \text{subject to} \quad & u \geq 0 \end{aligned}$$

## KKT Conditions for QP

$$\begin{aligned} Hz + q + G'u &= 0 \\ u_i(G_i z - w_i) &= 0 \quad i = 1, \dots, m \\ u &\geq 0 \\ Gz - w &\leq 0 \end{aligned}$$

## Active Constraints and Degeneracies

### Definition 2.4

The QP is said to be primal degenerate if there exists a  $z^* \in Z^*$  such that LICQ does not hold true at  $z^*$

### Definition 2.5

The QP is said to be dual degenerate if the dual problem is primal degenerate

**It can be noted that from the dual QP, shown above, that all the constraints are independent. Therefore, LICQ always holds for dual QPs and dual degeneracy can never occur for QPs with  $H \succ 0$**

## Mixed-Integer Optimization

If the decision set  $Z$  is the Cartesian product of a binary set and a real Euclidean space ( $Z \subseteq \{[z_c, z_b]: z_c \in \mathbb{R}^{s_c}, z_b \in \{0, 1\}^{s_b}\}$ ), then the optimization problem is said to be *mixed integer*.

This problem can be represented by:

$$\begin{aligned} \inf_{[z_c, z_b]} \quad & 1/2z'H z + q'z + r \\ \text{subject to} \quad & G_c z_c + G_b z_b \leq w \\ & z_c \in \mathbb{R}^{s_c}, z_b \in \{0, 1\}^{s_b} \\ & z = [z_c, z_b] \end{aligned}$$

Five cases can occur if the solution set  $P_{z_b}$  is not empty:

1. The solution is unbounded
2. The solution is bounded
3. The solution is bounded and there are infinitely many optimizers corresponding to the same integer value. This cannot happen if  $H_c \succ 0$
4. The solution is bounded and there are finite many optimizers corresponding to different integer values
5. The union of case 3 and 4

## Numerical Methods for Optimization

### Unconstrained Optimization (Numerical Methods 2.0)

The unconstrained optimization problem for which we minimize (or maximize) some function  $f(z)$  is interested in areas where the functions are continuously differentiable (think gradient). Via this method, we focus on descent methods. These methods obtain the next *iterate*  $z^{k+1}$  from the current iterate  $z^k$  by taking some step  $h^k$  along a certain direction  $d^k$ .

$$z^{k+1} = z^k + h^k d^k$$

## Gradient Methods

### Classic Gradient Method

**Lemma 3.1** (L-Smoothness: Second-Order Characterization):

Let  $f$  be twice continuously differentiable. Then the gradient of  $f$  is Lipschitz continuous with Lipschitz constant  $L > 0$  iff for all  $z \in \mathbb{R}^s$

$$\|\nabla f(z)\| \leq L$$

Recall that the gradient of  $f$  returns the direction of steepest local ascent. Therefore, by taking the “anti-gradient” direction, we find the direction of local steepest descent.

---

**Algorithm 1:** Gradient method for smooth convex optimization

---

**Input** : Initial iterate  $z^0 \in \mathbb{R}^s$ , Lipschitz constant  $L$  of  $\nabla f$

**Output:** Point close to  $z^*$

```
1 repeat
2   |  $z^{k+1} \leftarrow z^k - \frac{1}{L} \nabla f(z^k)$ ;
3 until Stopping criterion satisfied;
```

---

Where  $k^k = 1/L$ .

**Fast Gradient Method** The idea is to take the gradient at an affine combination of the previous two iterates

$$t^k = z^k + \beta^{k-1}(z^k - z^{k-1})$$

For some carefully chosen  $\beta^{k-1} \in (0, 1)$

---

**Algorithm 2:** Fast gradient method for smooth convex optimization

---

**Input** : Initial iterates  $z^0 \in \mathbb{R}^s$ ,  $y^0 = z^0$ ,  $\alpha^0 = 1/2(\sqrt{5} - 1)$ , Lipschitz constants of  $L$  of  $\nabla f$ .

**Output:** Point close to  $z^*$

```
1 repeat
2   |  $z^{k+1} \leftarrow z^k - \frac{1}{L} \nabla f(z^k)$  ;
3   |  $\alpha^{k+1} \leftarrow \frac{\alpha^k}{2}(\sqrt{\alpha^{k2} + 4} - \alpha^k)$  ;
4   |  $\beta \leftarrow \frac{\alpha^k(1-\alpha^k)}{\alpha^{k2} + \alpha^{k+1}}$  ;
5   |  $y^{k+1} \leftarrow z^{k+1} + \beta^k(z^{k+1} - z^k)$ ;
6 until Stopping criterion satisfied;
```

---

**Strong Convex Problem** In the case where the objective function is strongly convex, linear convergence rates can be established for both the gradient method and the fast gradient method.

**Preconditioning** In the case of strongly convex problems, the number of iterations required to find a minimizer depends on the condition number  $\kappa_f$ . High condition values usually equate to high number of iterations to converge. To improve the conditioning of the problem, one can solve an equivalent problem by a variable transformation:

$$y = Pz$$

with  $P$  invertible. The aim is to choose a preconditioner  $P$  such that the new objective function  $h(y) = f(P^{-1}z)$  has a smaller condition number.

### Newton's Method

In the following, we assume  $f$  to be  $\mu$ -strongly convex; hence the Hessian  $\nabla^2 f$  is positive definite.

---

**Algorithm 3:** Newton's method for minimizing  $f$  (strongly convex case)

---

**Input** : Initial iterates  $z^0 \in \mathbb{R}^s$

**Output:** Point close to  $z^*$

```
1 repeat
2   | Newton direction:  $d^k \leftarrow [\nabla^2 f(z^k)]^{-1} \nabla f(z^k)$  ;
3   | Line search: find  $h^k > 0$  such that  $f(z^k + h^k d^k) < f(z^k)$ ;
4   |  $z^{k+1} \leftarrow z^k + h^k d^k$ ;
5 until Stopping criterion satisfied;
```

---

The idea of Newton's method is to approximate a function by a local quadratic model based on the Taylor expansion. Although Newton's method is one of the most powerful algorithms for unconstrained optimization for twice differentiable function, the disadvantage is that you have to calculate the Hessian which can be numerically ill-conditioned which may make the cost of solving the linear system in the Newton's method much more expensive than the gradient method.

**Preconditioning and Affine Invariance** Preconditioning has no effect for Newton's method.

**Newton's Method with Linear Equality Constraints**

$$\begin{bmatrix} \nabla^2 f(z) & A' \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ y \end{bmatrix} = \begin{bmatrix} -\nabla f(z) \\ 0 \end{bmatrix}$$

Where  $y$  are the Lagrange multipliers associated with the equality constraints.

## Line Search Methods

A line search method determines the step length  $h^k$  that is taken from the current iterate along a search direction.

**Backtracking Line Search for Newton's Method**

---

**Algorithm 4:** Backtracking Line Search for Newton's Method

---

**Input** : Current iterate  $z^k$ , descent direction  $\Delta z^k$ , constants  $c \in (0, 1)$  and  $t \in (0, 1)$

**Output:** Step size  $h^k$  ensuring Wolfe conditions

```

1  $h^k \leftarrow 1$ ;
2 repeat
3    $h^k \leftarrow th^k$ ;
4 until Armijo's condition holds ( $f(z^k + h^k \Delta z^k) \leq F(z^k) + ch^k \Delta f(z^k)' \Delta z^k$ ;
```

---

## Constrained Optimization

### Gradient Projection Method

The solution minimization problem can be defined as:

$$k^{z+1} = \pi_S(z^k - 1/L \nabla f(z^k))$$

where

$$\pi_S(z) = \arg \min_{y \in S} 1/2 \|y - z\|^2$$

Where  $\pi_S(z)$  returns the value along the domain (just think the  $x$  value in  $f(x)$ ) that minimizes the  $z$ . If that doesn't make sense, try looking here.

---

**Algorithm 5:** Gradient method for smooth convex optimization

---

**Input** : Initial iterate  $z^0 \in \mathbb{R}^s$ , Lipschitz constant  $L$  of  $\nabla f$

**Output:** Point close to  $z^*$

```

1 repeat
2    $z^{k+1} \leftarrow \pi_S(z^k - \frac{1}{L} \nabla f(z^k))$ ;
3 until Stopping criterion satisfied;
```

---



---

**Algorithm 6:** Fast gradient method for smooth convex optimization

---

**Input** : Initial iterates  $z^0 \in \mathbb{R}^s$ ,  $y^0 = z^0$ ,  $\alpha^0 = 1/2(\sqrt{5} - 1)$ , Lipschitz constants of  $L$  of  $\nabla f$ .

**Output:** Point close to  $z^*$

```

1 repeat
2    $z^{k+1} \leftarrow \pi_S(z^k - \frac{1}{L} \nabla f(z^k))$ ;
3    $\alpha^{k+1} \leftarrow \frac{\alpha^k}{2} (\sqrt{\alpha^{k2} + 4} - \alpha^k)$ ;
4    $\beta \leftarrow \frac{\alpha^k(1-\alpha^k)}{\alpha^{k2} + \alpha^{k+1}}$ ;
5    $y^{k+1} \leftarrow z^{k+1} + \beta^k (z^{k+1} - z^k)$ ;
6 until Stopping criterion satisfied;
```

---

When the projection operator for set  $S$  can be evaluated efficiently, gradient projection methods have been shown to work very well in practice (see Table 3.1 for a list of convex sets that  $S$  can be found).



**Solution in Dual Domain** If a given set  $S$  cannot be calculated efficiently, one can still use the dual approach to solve the given problem. As we have seen before, we write the Lagrangian

$$\begin{aligned} d^* &= \max_v d(v) \\ \text{where} \\ d(v) &= \min_{z \in K} f(z) + v'(Az - b) \end{aligned}$$

## Interior Point Methods

The Primary Interior Point Method problem can be represented as:

$$\begin{aligned} \min_z \quad & f(z) \\ \text{subject to} \quad & Az = b \\ & g(z) \leq 0 \end{aligned}$$

**Primal Barrier Methods** The idea is to convert the constrained optimization problem shown above into an unconstrained problem (with respect to inequalities) by means of a barrier function  $\Phi_g$

$$\begin{aligned} z^*(\mu) \in \arg \min \quad & f(z) + \mu \Phi_g(z) \\ \text{subject to} \quad & Az = b \end{aligned}$$

The process is as described in the following algorithm

---

**Algorithm 7:** Fast gradient method for smooth convex optimization

---

**Input :** Strictly feasible initial iterate  $z^0$  w.r.t  $g(z) \leq 0$ ,  $\mu^0$ ,  $\kappa > 1$  ;  
tolerance  $\epsilon > 0$

**Output:** Point close to  $z^*$

```

1 repeat
2   Compute  $z^*(\mu^k)$  by minimizing  $f(z) + \mu^k \Phi_g(z)$  subject to  $Az = b$  starting from the previous solution  $z^{k-1}$ 
   (usually by Newton's method) This is called 'centering step' ;
3   Update:  $z^k \leftarrow z^*(\mu^k)$ ;
4   Stopping criterion: Stop if  $m\mu^k < \epsilon$  ;
5   Decrease barrier parameter:  $\mu^{k+1} \leftarrow \mu^k / \kappa$  ;
6 until Stopping criterion satisfied;
```

---

**Logarithmic Barrier** To use the barrier function method with Newton's method, the barrier function must be twice differentiable. That is why a logarithmic function is a good candidate.

$$\Phi_g = - \sum_{i=1}^m \ln(-g_i(z))$$

**Primal-Dual Methods** The general idea of the primal-dual interior point method is to solve the KKT conditions by a modified version of Newton's Method. This problem is represented as:

$$\begin{aligned} \nabla f(z) + A'z + G(z)'u &= 0 \\ Az - b &= 0 \\ g(z) + s &= 0 \\ s_i u_i &= 0 \quad i = 1, \dots, m \\ (s, u) &\geq 0 \end{aligned}$$

Where  $s$  denotes slack variables for the inequality constraints.

## Central Path

$$\begin{aligned} \nabla f(z) + A'z + G(z)'u &= 0 \\ Az - b &= 0 \\ g(z) + s &= 0 \\ s_i u_i &= \tau \quad i = 1, \dots, m, \tau > 0 \\ (s, u) &\geq 0 \end{aligned}$$

**Newton Directions and Centering** A search direction is taken by linearizing the central path problem above at the current iterate and solving. This method results in a vector  $\nu$  that allows for a corrective term called “centering”. In conducting with a weighing term  $\sigma$  known as the centering parameter, you can control how much the weight the solution for the Newton’s Method results are versus the centering.

### Active Set Method

The aim of the method is to identify the set of constraints at the solution. Once the set is known, the solution can be easily computed.

**Active Set Method for LP (Simplex Method)** Central to active set methods for LP is the observation that a solution is always attained at a vertex of the polyhedral feasible set. This fact is referred to the *fundamental theorem of linear programming* and gives reasoning to the methods’ alternative naming *simplex methods*.

For initialization, the simplex method requires vertex and an associated set of linearly independent constraint vectors.