

QP/LP Solvers

Barrier Interior-point Method

Input: strictly feasible z , $\kappa := \kappa^{(0)}$, $0 < \mu < 1$, tolerance $\epsilon > 0$

$\min_z \{f(z) \mid g(z) \leq 0\}$

- 1. **Centering step:** Compute $z^*(\kappa)$ by minimizing $f(z) + \kappa\phi(z)$ starting from z
- 2. **Update** $z := z^*(\kappa)$
- 3. **Stopping criterion:** Stop if $m\kappa < \epsilon$
- 4. **Decrease barrier parameter:** $\kappa := \mu\kappa$

• Gradient descent: $\Delta z := -\nabla f(z)$

$$\nabla f(z)^T \Delta z = -\nabla f(z)^T \nabla f(z) = -\|\nabla f(z)\| < 0$$

• Best: Newton method $\Delta z = -\nabla^2 f(z)^{-1} \nabla f(z)$

Newton step for Quadratic Programming

$$\min_z \left\{ \frac{1}{2} z^T H z \mid G z \leq d \right\}$$

• Barrier method:

$$\min_z f(z) + \kappa\phi(z) = \min_z \frac{1}{2} z^T H z - \kappa \sum_{i=1}^m \log(g_i - g_i z)$$

where g_1, \dots, g_m are the rows of G .

• The gradient and Hessian of the barrier function are:

$$\nabla\phi(z) = \sum_{i=1}^m \frac{1}{d_i - g_i z} g_i^T, \nabla^2\phi(z) = \sum_{i=1}^m \frac{1}{(d_i - g_i z)^2} g_i^T g_i$$

• Newton step: $(\nabla^2 f(z) + \kappa \nabla^2 \phi(z)) \Delta z_{nt} = -\nabla f(z) - \kappa \nabla \phi(z)$

$$(H + \kappa \sum_{i=1}^m \frac{1}{(d_i - g_i z)^2} g_i^T g_i) \Delta z_{nt} = -Hz - \kappa \sum_{i=1}^m \frac{1}{d_i - g_i z} g_i^T$$

Logarithmic Barrier Function

$$\phi(z) = -\sum_{i=1}^m \log(-g_i(z)), \quad \text{domain } \phi = \{z \mid g_1(z) \leq 0, \dots, g_m(z) \leq 0\}$$

$$\nabla\phi(z) = \sum_{i=1}^m \frac{1}{-g_i(z)} \nabla g_i(z)$$

$$\nabla^2\phi(z) = \sum_{i=1}^m \frac{1}{g_i(z)^2} \nabla g_i(z) \nabla g_i(z)^T + \frac{1}{-g_i(z)} \nabla^2 g_i(z)$$

descent direction : $\nabla f(z)^T \Delta z_{nt} = -\nabla f(z)^T \nabla^2 f(z)^{-1} \nabla(z) < 0$

Lyapunov equations

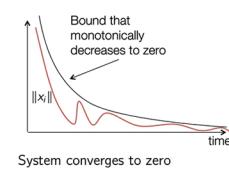
We assume $A \in \mathbb{R}^{n \times n}$, $P = P^T \in \mathbb{R}^{n \times n}$. It follows that $Q = Q^T \in \mathbb{R}^{n \times n}$.

Discrete-time linear system: for $x(t+1) = Ax(t)$, $V(z) = z^T P z$, we have $\Delta V(z) = -z^T Q z$, where P, Q satisfy (discrete-time) Lyapunov equation $A^T P A - P + Q = 0$.

Lyapunov theorems

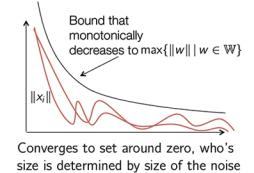
- If $P > 0$, $Q > 0$, then system is (globally asymptotically) stable.
- If $P > 0$, $Q \geq 0$, and (Q, A) observable, then system is (globally asymptotically) stable.
- If $P > 0$, $Q \geq 0$, then all trajectories of the system are bounded
- If $Q \geq 0$, then the sublevel sets $\{z \mid z^T P z \leq a\}$ are invariant. (These are ellipsoids if $P > 0$)
- If $P \geq 0$ and $Q \geq 0$, then A is not stable.

Asymptotic stability



System converges to zero

ISS stability



Converges to set around zero, who's size is determined by size of the noise

Invariant sets :

Positive Invariant set

A set \mathcal{O} is said to be a positive invariant set for the autonomous system $x_{i+1} = f(x_i)$ if

$$x_i \in \mathcal{O} \Rightarrow x_i \in \mathcal{O}, \quad \forall i \in \{0, 1, \dots\}$$

Pre Set

Given a set S and the dynamic system $x^+ = f(x)$, the pre-set of S is the set of states that evolve into the target set S in one time step:

$$\text{pre}(S) := \{x \mid f(x) \in S\}$$

Theorem: Geometric condition for invariance

A set \mathcal{O} is a positive invariant set if and only if

$$\mathcal{O} \subseteq \text{pre}(\mathcal{O})$$

Polyhedron

A polyhedron is the intersection of a finite number of halfspaces.

$$P := \{x \mid a_i^T x \leq b_i, \quad i = 1, \dots, n\}$$

A polytope is a bounded polyhedron.

Conceptual Algorithm to Compute Invariant Set

Input: f, \mathbb{X}

Output: \mathcal{O}_∞

```

 $\Omega_0 \leftarrow \mathbb{X}$ 
loop
   $\Omega_{i+1} \leftarrow \text{pre}(\Omega_i) \cap \Omega_i$ 
  if  $\Omega_{i+1} = \Omega_i$  then
    return  $\mathcal{O}_\infty = \Omega_i$ 
  end if
end loop

```

The algorithm generates the set sequence $\{\Omega_i\}$ satisfying $\Omega_{i+1} \subseteq \Omega_i$ for all $i \in \mathbb{N}$ and it terminates when $\Omega_{i+1} = \Omega_i$, so that Ω_i is the maximal positive invariant set \mathcal{O}_∞ for $x^+ = f(x)$.

Convex hull

For any subset S of \mathbb{R}^d , the convex hull $\text{conv}(S)$ of S is the intersection of all convex sets containing S . Since the intersection of two convex sets is convex, it is the smallest convex set containing S .

Proposition: Convex hull

The convex hull of a set $S \subseteq \mathbb{R}^d$ is

Given a set of points $\{v_1, \dots, v_k\}$ in \mathbb{R}^d , their convex hull is

$$\text{conv}(\{v_1, \dots, v_k\}) := \left\{ x \mid x = \sum_i \lambda_i v_i, \quad \lambda_i \geq 0, \quad \sum_i \lambda_i = 1 \quad \forall i = 1, \dots, k \right\}$$

Intersection

The intersection $I \subseteq \mathbb{R}^n$ of sets $S \subseteq \mathbb{R}^n$ and $T \subseteq \mathbb{R}^n$ is

$$I = S \cap T := \{x \mid x \in S \text{ and } x \in T\}$$

Intersection of polytopes in inequality form is easy:

$$S := \{x \mid Cx \leq c\} \quad T := \{x \mid Dx \leq d\} \quad S \cap T = \left\{ x \mid \begin{bmatrix} C \\ D \end{bmatrix} x \leq \begin{bmatrix} c \\ d \end{bmatrix} \right\}$$

Pre-Set Computation: Controlled System

Consider the system $x^+ = Ax + Bu$ under the constraints $u \in \mathbb{U} := \{u \mid Gu \leq g\}$ and the set $S := \{x \mid Fx \leq f\}$.

$$\text{pre}(S) := \{x \mid \exists u \in \mathbb{U} \text{ s.t. } f(x, u) \in S\}$$

$$= \{x \mid \exists u \in \mathbb{U}, Ax + Bu \in S\}$$

$$= \left\{ x \mid \exists u, \begin{bmatrix} FA & FB \\ 0 & G \end{bmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \leq \begin{bmatrix} f \\ g \end{bmatrix} \right\}$$

Polytopic Projection

Given a polytope $P = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^d \mid Cx + Dy \leq b\}$, find a matrix E and vector e , such that the polytope

$$P_\pi = \{x \mid Ex \leq e\} = \{x \mid \exists y, (x, y) \in P\}$$

Robust MPC:

Robust Pre Set

Given a set Ω and the dynamic system $x^+ = f(x, w)$, the pre-set of Ω is the set of states that evolve into the target set Ω in one time step for all values of the disturbance $w \in \mathbb{W}$:

$$\text{pre}^\mathbb{W}(\Omega) := \{x \mid f(x, w) \in \Omega \text{ for all } w \in \mathbb{W}\}$$

Computing Robust Pre-Sets for Linear Systems

Goal: Given the system $f(x, w) = Ax + w$, and the set $\Omega := \{x \mid Fx \leq f\}$, compute $\text{pre}^\mathbb{W}(\Omega)$.

$$\text{pre}^\mathbb{W}(\Omega) = \{x \mid Ax + w \in \Omega, \quad \forall w \in \mathbb{W}\} = \{x \mid FAx + Fw \leq f, \quad \forall w \in \mathbb{W}\}$$

$$\text{pre}^\mathbb{W}(\Omega) = \left\{ x \mid \begin{array}{l} FAx \leq f - \max_{w \in \mathbb{W}} Fw \end{array} \right\} = \{x \mid FAx \leq f - h_w(F) = A(\Omega \ominus \mathbb{W})\}$$

Computing Robust Invariant Sets

...

Tube-MPC Assumptions

Much the same as for nominal MPC:

1. The stage cost is a positive definite function, i.e. it is strictly positive and only zero at the origin
2. The terminal set is invariant for the nominal system under the local control law $\kappa_f(z)$:
$$z^+ = Az + B\kappa_f(z) \in \mathcal{X}_f \quad \text{for all } z \in \mathcal{X}_f$$

All tightened state and input constraints are satisfied in \mathcal{X}_f :
$$\mathcal{X}_f \subseteq \mathbb{X} \ominus \mathcal{E}, \quad \kappa_f(z) \in \mathbb{U} \ominus \mathcal{E} \quad \text{for all } z \in \mathcal{X}_f$$
3. Terminal cost is a continuous Lyapunov function in the terminal set \mathcal{X}_f :
$$V_f(Az + B\kappa_f(z)) - V_f(z) \leq -l(z, \kappa_f(z)) \quad \text{for all } z \in \mathcal{X}_f$$

NMPC

Runge Kutta 2

Consider the ODE $\dot{x} = f(x)$

Given $x = x(t)$, we want to compute $x^+ = x(t+h)$

Compute a second-order Taylor series expansion

$$x^+ = x + h\dot{x} + \frac{h^2}{2} \ddot{x} + \mathcal{O}(h^3)$$

Take Jacobian of f to compute \ddot{x}

$$\ddot{x} = J_f(x)\dot{x} = J_f(x)f(x)$$

The Taylor series expansion is now

$$\begin{aligned} x^+ &= x + hf(x) + \frac{h^2}{2} J_f(x)f(x) + \mathcal{O}(h^3) \\ &= x + \frac{h}{2} f(x) + \frac{h}{2} (f(x) + hJ_f(x)f(x)) + \mathcal{O}(h^3) \end{aligned}$$

Runge-Kutta 4 - The most common version

Consider the time dependent ODE: $\dot{x} = f(t, x)$

$$x_{k+1} = x_k + h \left(\frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6} \right)$$

$$k_1 = f(t_k, x_k)$$

$$k_2 = f(t_k + \frac{h}{2}, x_k + \frac{h}{2}k_1)$$

$$k_3 = f(t_k + \frac{h}{2}, x_k + \frac{h}{2}k_2)$$

$$k_4 = f(t_k + h, x_k + h k_3)$$

Minkowski sum :

Given $P := \{x \mid Tx \leq t\}$ and $Q := \{x \mid Rx \leq r\}$, the Minkowski sum is:

$$P \oplus Q := \{x + y \mid x \in P, y \in Q\}$$

$$= \{z \mid \exists x, y \in P, Ry \leq z - Tx \leq t, Ry \leq r\}$$

$$= \left\{ z \mid \exists y, \begin{bmatrix} T & -R \\ 0 & I \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} t \\ r \end{pmatrix} \right\}$$

Subset test :

Problem: Is $P := \{x \mid Cx \leq c\}$ contained in $Q := \{x \mid Dx \leq d\}$?

The statement is true if $P \subset \{x \mid Di_x \leq d_i\}$ for each row D_i of D .

$$h_P(D_i) := \max_x D_i x$$

s.t. $Cx \leq c$



LQR Regulation

Linear Quadratic Regulator

$$V^*(x_0) := \min_u \sum_{k=0}^{\infty} x_k^T Q x_k + u_k^T R u_k \quad \text{s.t. } x_{k+1} = A x_k + B u_k$$

Can solve the infinite-horizon predictive control problem in closed-form:

$$P = Q + A^T P A - A^T P B (R + B^T P B)^{-1} B^T P A$$

The optimal input is the constant state feedback

$$u = K x$$

$$K = -(R + B^T P B)^{-1} B^T P A$$

The optimal cost function $V^*(x) = x^T P x$ is a Lyapunov function for the closed-loop system $x^+ = (A + BK)x$.

Lemma: Lyapunov function for LQR

The optimal value function $V^*(x) = x^T P x$ is a Lyapunov function for the system $x^+ = (A + BK)x$ where $K = -(R + B^T P B)^{-1} B^T P A$ and P solves

$$P = Q + A^T P A - A^T P B (R + B^T P B)^{-1} B^T P A$$

for some $Q \succeq 0, R > 0$.

Solving Infinite-Horizon LQR

Fact: $V^*(x)$ is quadratic, $V^*(x) = x^T P x$ for $P \succ 0^2$

Bellman equation:

$$V_i^*(x_i) := \min_u l(x_i, u_i) + V_{i+1}^*(Ax_i + Bu_i)$$

If $V_i^*(\cdot) = V_{i+1}^*(\cdot)$, then $V_j^*(\cdot) = V_{i+1}^*(\cdot)$ for all $j \leq i$.

minimizing gives $u^* = -(R + B^T P B)^{-1} B^T P A x$, giving

$$x^T P x = x^T Q x + u^T R u + (Ax + Bu)^T P (Ax + Bu)$$

$$x^T P x = x^T (Q + A^T P A - A^T P B (R + B^T P B)^{-1} B^T P A) x$$

Consider the DP iteration:

$$V_i^*(x_i) := \min_u l(x_i, u_i) + V_{i+1}^*(Ax_i + Bu_i)$$

If $V_i^*(\cdot) = V_{i+1}^*(\cdot)$, then $V_j^*(\cdot) = V_{i+1}^*(\cdot)$ for all $j \leq i$.

Dynamic Programming

1. Start at step N and compute $V_N^*(x_N) := \min_u x_N^T Q x_N + u_N^T R u_N$

$= x_N^T Q x_N \quad H_N := Q$

2. Iterate backwards for $i = N-1 \dots 0$ (DP iteration)

$$V_i^*(x_i) := \min_u x_i^T Q x_i + u_i^T R u_i + V_{i+1}^*(Ax_i + Bu_i)$$

$$u_i^*(x_i) = K_i x_i \quad K_i = -(R + B^T H_{i+1} B)^{-1} B^T H_{i+1}$$

$$V_i^*(x_i) = x_i^T H_i x_i \quad H_i := Q + K_i^T R K_i + (A + B K_i)^T H_{i+1} (A + B K_i)$$

3. $V^*(x_0) := V_0^*(x_0)$ and the optimal controller is the optimizer $u_0^*(x_0)$

MPC Properties

Lyapunov function

Let \mathcal{X} be a positively invariant set for system $x(k+1) = f_k(x(k))$ containing a neighborhood of the origin in its interior. A function $V : \mathcal{X} \rightarrow \mathbb{R}_+$ is called a **Lyapunov function** in \mathcal{X} if for all $x \in \mathcal{X}$:

$$\begin{aligned} V(x) &> 0 \forall x \neq 0, \quad V(0) = 0, \\ V(x(k+1)) - V(x(k)) &\leq 0 \end{aligned}$$



Theorem: (e.g., [Vidyasager, 1993])

If a system admits a Lyapunov function in \mathcal{X} , then the equilibrium point at the origin is (**Lyapunov**) stable in \mathcal{X} .

Definition: Feasible set

The **feasible set** \mathcal{X}_N is defined as the set of initial states x for which the MPC problem with horizon N is feasible, i.e.

$$\mathcal{X}_N := \{x \mid \exists [u_0, \dots, u_{N-1}] \text{ such that } Cu_i + Dx_i \leq b, i = 1, \dots, N\}$$

Linear MPC with Quadratic Cost

Standard formulation:

- Quadratic performance measure
- Linear system dynamics
- \mathcal{X}_f and \mathcal{U} are polyhedra

$$\begin{aligned} \min_u & \sum_{i=0}^{N-1} x_i^T Q x_i + u_i^T R u_i + x_N^T Q_f x_N \\ \text{s.t. } & x_i \in \mathbb{X} \quad i \in \{1, \dots, N-1\} \\ & u_i \in \mathbb{U} \quad i \in \{0, \dots, N-1\} \\ & x_N \in \mathcal{X}_f \\ & x_{i+1} = Ax_i + Bu_i \end{aligned}$$

Assumptions: $Q = Q^T \succeq 0, Q_f = Q_f^T \succ 0, R = R^T \succ 0$

Delta-Formulation for tracking

Define deviation variables that (in the linear case) satisfy the same model equations:

$$\begin{aligned} \Delta x = x - x_s & \Rightarrow \Delta x_{k+1} = x_{k+1} - x_s \\ \Delta u = u - u_s & \Rightarrow = Ax_k + Bu_k - (Ax_s + Bu_s) \\ & = A\Delta x_k + B\Delta u_k \end{aligned}$$

Constraints for deviation variables:

$$\begin{aligned} H_x x \leq k_x & \Rightarrow H_x \Delta x \leq k_x - H_x x_s \\ H_u \leq k_u & \Rightarrow H_u \Delta u \leq k_u - H_u u_s \end{aligned}$$

Apply regulation problem to new system in Delta-Formulation:

$$\begin{aligned} \min_u & \sum_{i=0}^{N-1} \Delta x_i^T Q \Delta x_i + \Delta u_i^T R \Delta u_i + V_r(\Delta x_N) \\ \text{s.t. } & \Delta x_0 = \Delta x \\ & \Delta x_{i+1} = A\Delta x_i + B\Delta u_i \\ & H_x \Delta x_i \leq k_x - H_x x_s \\ & H_u \Delta u_i \leq k_u - H_u u_s \\ & \Delta x_N \in \mathcal{X}_f \end{aligned}$$

pLCPs

Parametric Linear Complementarity Problem

Given matrices M, q and Q , find functions $w(x), z(x)$ such that

$$\begin{aligned} w - Mz &= q + Qx \\ w^T z &= 0 \\ w, z &\geq 0 \end{aligned}$$

Definition 9.1. KKT Necessary and Sufficient Optimality Conditions

Consider the following problem: $\min f(x)$,

$$Ax \leq b,$$

1. **Stationarity**: $\nabla f(x^*) + A^T \lambda = 0$

Gradient is in the normal cone.

2. **Primal Feasibility**: $Ax^* \leq b$

Ensures that x^* is in the feasible polytope.

3. **Dual Feasibility**: $\lambda \geq 0$

Ensures that Stationarity is enforced.

4. **Complementarity**: $\lambda^T (Ax - b) = 0$

Normal cone contains only active constraints. Same as saying that either $\lambda_i = 0$ or $a_i x = b_i, \forall i$

Parametric quadratic optimization problem:

$$\begin{aligned} J^*(x) := \min_u & \frac{1}{2} u^T Qu + (Fx + f)^T u \\ \text{s.t. } & Gu \geq Ex + e \\ & u \geq 0 \end{aligned}$$

KKT Conditions:

$$\Rightarrow Qu + Fx + f - G^T \lambda - \nu = 0 \quad \text{Stationarity}$$

$$\Rightarrow -s + Gu = Ex + e, \quad u \geq 0 \quad \text{Primary feasibility}$$

$$\lambda, \nu \geq 0 \quad \text{Dual feasibility}$$

$$\nu^T u = 0, \quad \lambda^T s = 0 \quad \text{Complementarity}$$

Stationarity

$$\Rightarrow \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{pmatrix} \nu \\ s \end{pmatrix} - \begin{bmatrix} Q & -G^T \\ G & 0 \end{bmatrix} \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{bmatrix} F \\ -E \end{bmatrix} x + \begin{bmatrix} f \\ -e \end{bmatrix}$$

Primal feasibility

$$\nu, s, u, \lambda \geq 0$$

Primal and dual feasibility

$$\nu^T u = s^T \lambda = 0$$

Complementarity

Standard pLCP:

$$Iw - Mz = Qx + q$$

$$w, z \geq 0, \quad w^T z = 0$$

Note: Optimizer is linear transform of pLCP solution $u^*(x) = [I \ 0] z(x)$

Lyapunov Function for LQR-Controlled System

Lemma: Lyapunov function for LQR

The optimal value function $V^*(x) = x^T P x$ is a Lyapunov function for the system $x^+ = (A + BK)x$ where $K = -(R + B^T P B)^{-1} B^T P A$ and P solves

$$P = Q + A^T P A - A^T P B (R + B^T P B)^{-1} B^T P A$$

for some $Q \succeq 0, R > 0$.

Dynamic Programming

1. Start at step N and compute $V_N^*(x_N) := \min_u x_N^T Q x_N + u_N^T R u_N$

$= x_N^T Q x_N \quad H_N := Q$

2. Iterate backwards for $i = N-1 \dots 0$ (DP iteration)

$$V_i^*(x_i) := \min_u x_i^T Q x_i + u_i^T R u_i + V_{i+1}^*(Ax_i + Bu_i)$$

$$u_i^*(x_i) = K_i x_i \quad K_i = -(R + B^T H_{i+1} B)^{-1} B^T H_{i+1}$$

$$V_i^*(x_i) = x_i^T H_i x_i \quad H_i := Q + K_i^T R K_i + (A + B K_i)^T H_{i+1} (A + B K_i)$$

3. $V^*(x_0) := V_0^*(x_0)$ and the optimal controller is the optimizer $u_0^*(x_0)$



Formalize Goals: Feasibility and Stability

Goal 1: Feasibility for all time

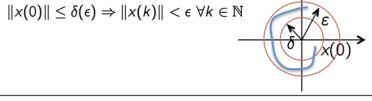
Definition: Recursive feasibility

The MPC problem is called **recursively feasible**, if for all feasible initial states feasibility is guaranteed at every state along the closed-loop trajectory.

Goal 2: Stability

Definition: Lyapunov stability

The equilibrium point at the origin of system $x_{k+1} = Ax_k + B\kappa(x_k) = f_k(x_k)$ is said to be (**Lyapunov**) **stable** in \mathcal{X} if for every $\epsilon > 0$, there exists a $\delta(\epsilon) > 0$ such that, for every $x(0) \in \mathcal{X}$:



Stability of MPC - Main Result

If we can choose/find an \mathcal{X}_f , κ_f , V_f and l such that:

1. The stage cost is a positive definite function, i.e. it is strictly positive and only zero at the origin
2. The terminal set is **invariant** under the local control law $\kappa_f(x)$:

$$V_f(x^+) - V_f(x) \leq -l(x, \kappa_f(x)) \text{ for all } x \in \mathcal{X}_f$$

MPC problem for tracking

Convergence

Assume target is feasible with $x_s \in \mathbb{X}, u_s \in \mathbb{U}$ and choose terminal weight $V_f(x)$ and constraint \mathcal{X}_f as in the regulation case satisfying:

- $\mathcal{X}_f \subseteq \mathbb{X}, Kx \in \mathbb{U}$ for all $x \in \mathcal{X}_f$
- $V_f(x^+) - V_f(x) \leq -l(x, Kx)$ for all $x \in \mathcal{X}_f$

If in addition the target reference x_s, u_s is such that

- $x_s \oplus \mathcal{X}_f \subseteq \mathbb{X}, K\Delta x + u_s \in \mathbb{U}$ for all $\Delta x \in \mathcal{X}_f$

then the closed-loop system converges to the target reference, i.e. $x_k \rightarrow x_s$ and therefore $y_k = Cx_k \rightarrow r$ for $k \rightarrow \infty$

Soft constrained MPC problem setup

$$\min_u \sum_{i=0}^{N-1} x_i^T Q x_i + u_i^T R u_i + \rho(\epsilon_i) + x_N^T P x_N + \rho(\epsilon_N)$$

$$\text{s.t. } x_{i+1} = Ax_i + Bu_i$$

$$H_x x_i \leq k_x + \epsilon_i,$$

$$H_u u_i \leq k_u,$$

$$\epsilon_i \geq 0$$

• Relax state constraints by introducing so called

slack variables $\epsilon_i \in \mathbb{R}^p$

- Quadratic penalty: $\rho(\epsilon_i) = \epsilon_i^T S \epsilon_i$
- Quadratic and linear norm penalty: $\rho(\epsilon_i) = \epsilon_i^T S \epsilon_i + s \|\epsilon_i\|_1$

Geometric Series

$$\sum_{n=0}^{N-1} q^n = \frac{1-q^N}{1-q} \quad \sum q^n = \frac{1}{1-q}$$

Algebraic Identities

$$(AB)^{-1} = B^{-1} A^{-1}$$

$$(ABC\dots)^{-1} = \dots C^{-1} B^{-1} A^{-1}$$

$$(A^T)^{-1} = (A^{-1})^T$$

$$(A+B)^T = A^T + B^T$$

$$(AB)^T = B^T A^T$$

$$(ABC\dots)^T = \dots C^T B^T A^T$$

Hessian Matrix

$$H_{ij}(f) = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Or compute each column by derivating the gradient for each variable