Function-Space Approach for Gradient Descent in Optimal Control

Maurice Filo

University of California, Santa Barbara
https://engineering.ucsb.edu/~filo/
Advisor: Bassam Bamieh

June 28, 2018

Overview

Problem Formulation: Function Space Approach

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T x^*(t)Qx(t) + u^*(t)Ru(t) \;\; dt$$

$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

Problem Formulation: Function Space Approach

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T x^*(t)Qx(t) + u^*(t)Ru(t) \quad dt$$
$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

Define:

$$z := \begin{bmatrix} x \\ u \end{bmatrix}; \qquad x = \mathcal{H}(u)$$

Problem Formulation: Function Space Approach

$$\begin{aligned} \underset{x,u}{\text{minimize}} \quad & J(x,u) = \frac{1}{2}\int_0^T x^*(t)Qx(t) + u^*(t)Ru(t) \quad dt \\ \text{subject to} \quad & \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0 \end{aligned}$$

Define:

$$z := \begin{bmatrix} x \\ u \end{bmatrix}; \qquad x = \mathcal{H}(u)$$

Then:

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & J(z) = \frac{1}{2}\langle z, Hz \rangle \qquad\qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \\ \text{subject to} \quad & x = \mathcal{H}(u) \end{aligned}$$

Problem Formulation: Function Space Approach

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T x^*(t)Qx(t) + u^*(t)Ru(t) \quad dt$$
$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

Define:

$$z := \begin{bmatrix} x \\ u \end{bmatrix}; \qquad x = \mathcal{H}(u)$$

Then:

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$
$$\text{subject to} \quad x = \mathcal{H}(u)$$

Unconstrained Optimization: $\mathcal{J}(u) := J(\mathcal{H}(u), u) = \frac{1}{2}\left\langle \begin{bmatrix} \mathcal{H}(u) \\ u \end{bmatrix}, H \begin{bmatrix} \mathcal{H}(u) \\ u \end{bmatrix} \right\rangle$

Problem Formulation: Function Space Approach

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T x^*(t)Qx(t) + u^*(t)Ru(t) \quad dt$$
$$\text{subject to} \quad \dot{x}(t) = f\big(x(t),u(t)\big); \quad x(0) = x_0$$

Define:

$$z := \begin{bmatrix} x \\ u \end{bmatrix}; \qquad x = \mathcal{H}(u)$$

Then:

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$
$$\text{subject to} \quad x = \mathcal{H}(u)$$

Unconstrained Optimization: $\mathcal{J}(u) := J(\mathcal{H}(u),u) = \frac{1}{2}\left\langle \begin{bmatrix} \mathcal{H}(u) \\ u \end{bmatrix}, H \begin{bmatrix} \mathcal{H}(u) \\ u \end{bmatrix} \right\rangle$

- First Order Method: Gradient Descent $\longrightarrow$ Cheap but Slow Convergence
- Second Order Method: Newton $\longrightarrow$ Fast Convergence but Expensive

Proposed Method

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T x^*(t)Qx(t) + u^*(t)Ru(t) \;\; dt$$
$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

Proposed Method: Keep cost functional & Dynamics separate!

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad\qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$
$$\text{subject to} \quad x = \mathcal{H}(u)$$

Proposed Method

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T x^*(t)Qx(t) + u^*(t)Ru(t) \;\; dt$$
$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

Proposed Method: Keep cost functional & Dynamics separate!

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$
$$\text{subject to} \quad x = \mathcal{H}(u)$$

Dynamical Constraint Set (Trajectories Manifold): [1]

$$x = \mathcal{H}(u) \Longleftrightarrow z \in \mathcal{M} \qquad \mathcal{M} = \Big\{ z = (x,u) : x = \mathcal{H}(u) \Big\}$$

[1] J. Hauser and D. G. Meyer, "The trajectory manifold of a nonlinear control system," in Decision and Control, 1998. Proceedings of the 37th IEEE Conference on, vol. 1, pp. 1034–1039, IEEE, 1998.

Proposed Method

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T x^*(t)Qx(t) + u^*(t)Ru(t) \quad dt$$
$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

Proposed Method: Keep cost functional & Dynamics separate!

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$
$$\text{subject to} \quad z \in \mathcal{M}$$

Dynamical Constraint Set (Trajectories Manifold): [1]

$$x = \mathcal{H}(u) \iff z \in \mathcal{M} \qquad \mathcal{M} = \Big\{ z = (x,u) : x = \mathcal{H}(u) \Big\}$$

[1] J. Hauser and D. G. Meyer, "The trajectory manifold of a nonlinear control system," in Decision and Control, 1998. Proceedings of the 37th IEEE Conference on, vol. 1, pp. 1034–1039, IEEE, 1998.

Precondition Constrained-Gradient Descent (PCGD)

$$\begin{aligned} &\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz \rangle \\ &\text{subject to} \quad z \in \mathcal{M} \end{aligned}$$

Precondition Constrained-Gradient Descent (PCGD)

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & J(z) = \frac{1}{2}\langle z, Hz\rangle \\ \text{subject to} \quad & z \in \mathcal{M} \end{aligned}$$

Similar in spirit to a projection-based Newton method developed by J. Hauser [2]

[2] Hauser, J. (2002). A projection operator approach to the optimization of trajectory functionals. IFAC Proceedings Volumes, 35(1), 377-382.

Precondition Constrained-Gradient Descent (PCGD)

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & J(z) = \frac{1}{2}\langle z, Hz \rangle \\ \text{subject to} \quad & z \in \mathcal{M} \end{aligned}$$

Similar in spirit to a projection-based Newton method developed by J. Hauser [2]

Two Key ideas:

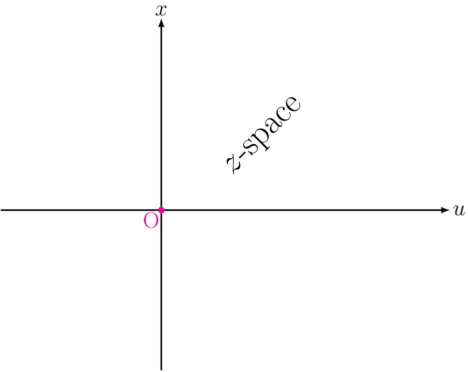
$\longrightarrow$ Two different types of projections
$\longrightarrow$ Preconditioning the state-control space (z-space)

[2] Hauser, J. (2002). A projection operator approach to the optimization of trajectory functionals. IFAC Proceedings Volumes, 35(1), 377-382.

Key Idea 1: Projections...

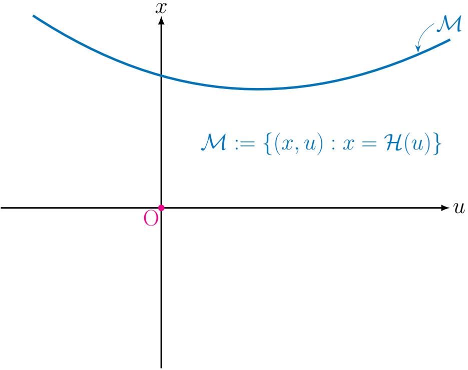
$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$



Key Idea 1: Projections...

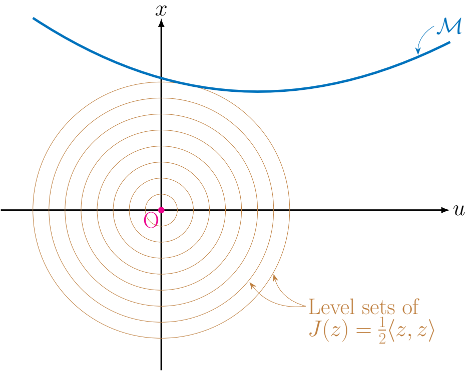
$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$



- $\mathcal{M}$: Dynamical Constraints Manifold

Key Idea 1: Projections...

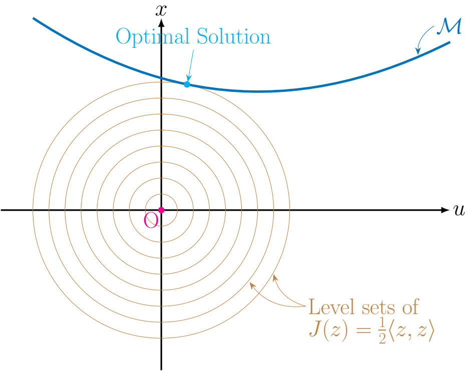
$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$



- $\mathcal{M}$: Dynamical Constraints Manifold
-
-
-
-
-

Key Idea 1: Projections...

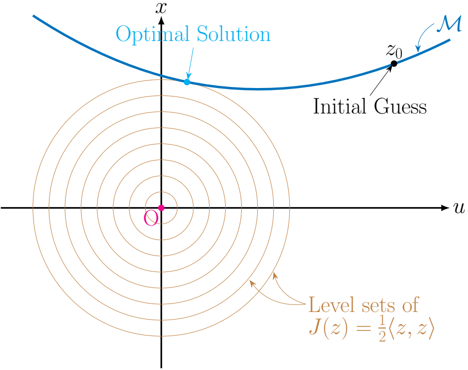
$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$



- $\mathcal{M}$: Dynamical Constraints Manifold

Key Idea 1: Projections...

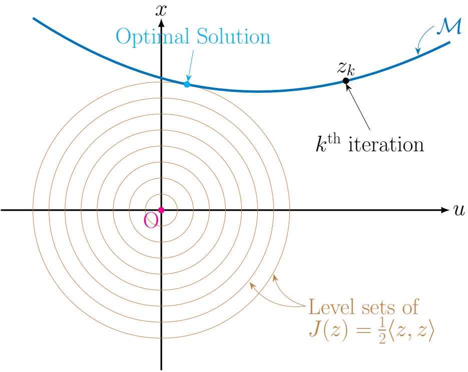
$$\begin{aligned} \underset{z}{\text{minimize}} \quad & J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I) \\ \text{subject to} \quad & z \in \mathcal{M} \end{aligned}$$



- $\mathcal{M}$: Dynamical Constraints Manifold

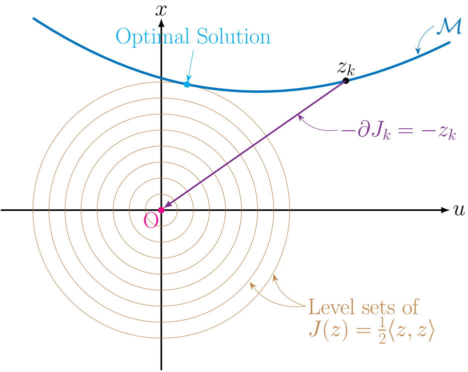
Key Idea 1: Projections...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$



- $\mathcal{M}$: Dynamical Constraints Manifold

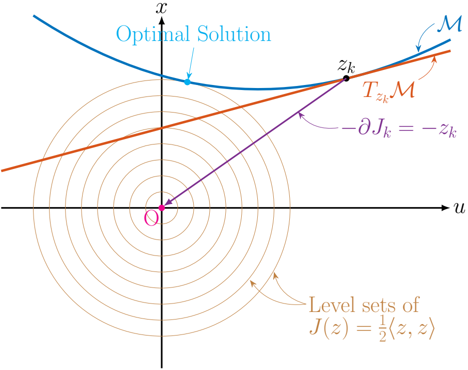
$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$



- $\mathcal{M}$: Dynamical Constraints Manifold
- ∂J_k: Gradient at Current Iteration

Key Idea 1: Projections...

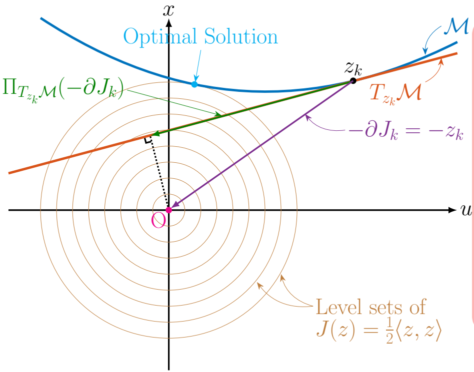
$$\begin{array}{ll} \underset{z}{\text{minimize}} & J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I) \\ \text{subject to} & z \in \mathcal{M} \end{array}$$



- $\mathcal{M}$: Dynamical Constraints Manifold
- ∂J_k: Gradient at Current Iteration
- $T_{z_k}\mathcal{M}$: Tangent Space of $\mathcal{M}$ at z_k

Key Idea 1: Projections...

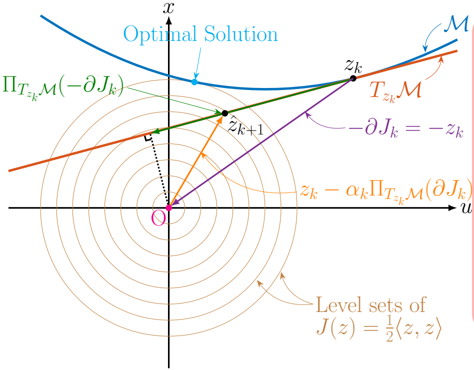
$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$



- $\mathcal{M}$: Dynamical Constraints Manifold
- ∂J_k: Gradient at Current Iteration
- $T_{z_k}\mathcal{M}$: Tangent Space of $\mathcal{M}$ at z_k
- $\Pi_{T_{z_k}\mathcal{M}}$: Linear Tangent Space Projection Operator

Key Idea 1: Projections...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$
$$\text{subject to} \quad z \in \mathcal{M}$$

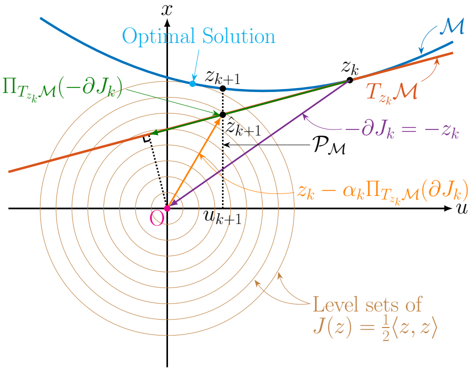


- $\mathcal{M}$: Dynamical Constraints Manifold
- ∂J_k: Gradient at Current Iteration
- $T_{z_k}\mathcal{M}$: Tangent Space of $\mathcal{M}$ at z_k
- $\Pi_{T_{z_k}\mathcal{M}}$: Linear Tangent Space Projection Operator
- α_k: Step size at current iteration
-

Key Idea 1: Projections...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$

$$\text{subject to} \quad z \in \mathcal{M}$$

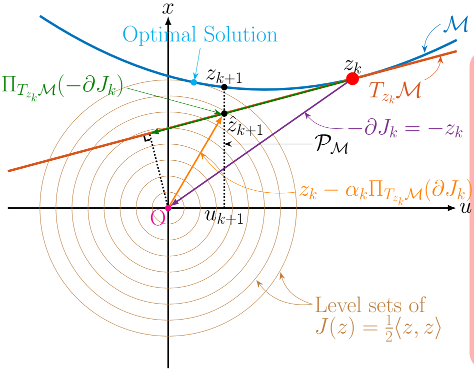


- $\mathcal{M}$: Dynamical Constraints Manifold
- ∂J_k: Gradient at Current Iteration
- $T_{z_k}\mathcal{M}$: Tangent Space of $\mathcal{M}$ at z_k
- $\Pi_{T_{z_k}\mathcal{M}}$: Linear Tangent Space Projection Operator
- α_k: Step size at current iteration
- $\mathcal{P}_\mathcal{M}$: Nonlinear Trajectory Projection Operator

Key Idea 1: Projections...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle \qquad (H = I)$$

$$\text{subject to} \quad z \in \mathcal{M}$$

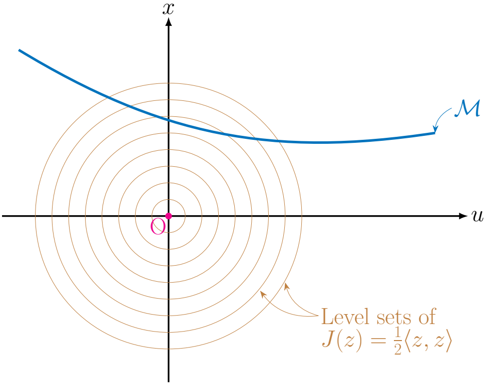


- $\mathcal{M}$: Dynamical Constraints Manifold
- ∂J_k: Gradient at Current Iteration
- $T_{z_k}\mathcal{M}$: Tangent Space of $\mathcal{M}$ at z_k
- $\Pi_{T_{z_k}\mathcal{M}}$: Linear Tangent Space Projection Operator
- α_k: Step size at current iteration
- $\mathcal{P}_{\mathcal{M}}$: Nonlinear Trajectory Projection Operator

For Spherical Level Sets: $\begin{cases} \hat{z}_{k+1} = z_k - \alpha_k \Pi_{T_{z_k}\mathcal{M}}(\partial J_k) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}}\left(\hat{z}_{k+1}\right) \end{cases}$

Special Case: Linear Dynamics

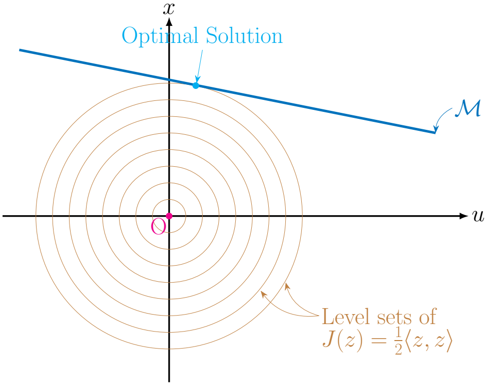
$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle$$
$$\text{subject to} \quad z \in \mathcal{M} \qquad \mathcal{M} := \{z = (x, u) : \dot{x} = Ax + Bu;\ x(0) = x_0\}$$



Special Case: Linear Dynamics

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle$$

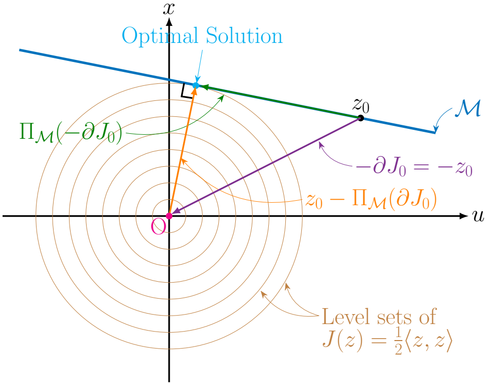
$$\text{subject to} \quad z \in \mathcal{M} \qquad \mathcal{M} := \{z = (x,u) : \dot{x} = Ax + Bu;\ x(0) = x_0\}$$



Special Case: Linear Dynamics

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, z\rangle$$

$$\text{subject to} \quad z \in \mathcal{M} \qquad \mathcal{M} := \{z = (x,u) : \dot{x} = Ax + Bu;\ x(0) = x_0\}$$

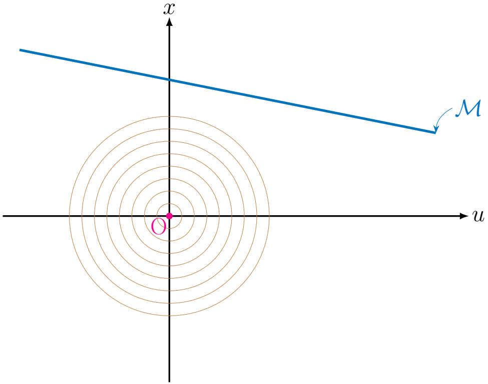


Converges in one iteration with step size $\alpha = 1$!

Special Case: Linear Dynamics

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad (H \neq I)$$

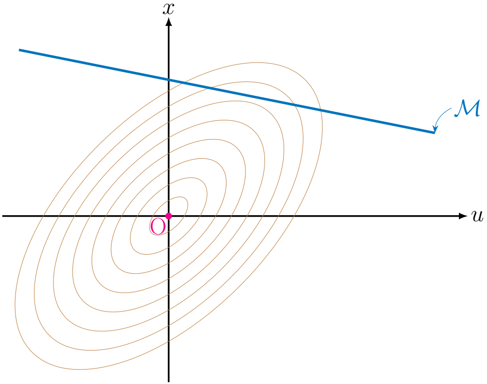
$$\text{subject to} \quad z \in \mathcal{M} \qquad \mathcal{M} := \{z = (x,u) : \dot{x} = Ax + Bu;\ x(0) = x_0\}$$



Special Case: Linear Dynamics

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz \rangle \qquad (H \neq I)$$

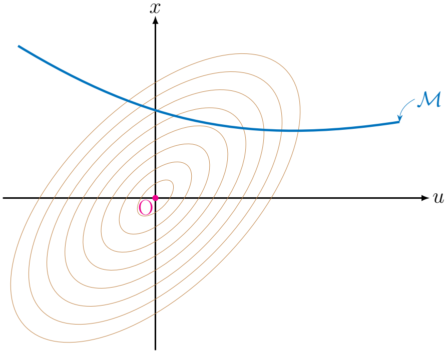
$$\text{subject to} \quad z \in \mathcal{M} \qquad \mathcal{M} := \{z = (x, u) : \dot{x} = Ax + Bu;\ x(0) = x_0\}$$



Ellipsoidal level sets: does not converge in one iteration!

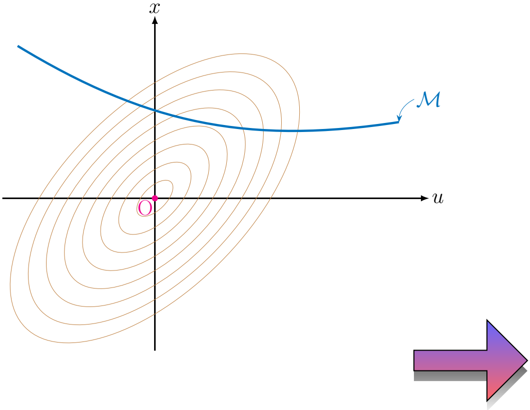
Key Idea 2: Preconditioning...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz \rangle$$
$$\text{subject to} \quad z \in \mathcal{M}$$



Key Idea 2: Preconditioning...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle$$
$$\text{subject to} \quad z \in \mathcal{M}$$

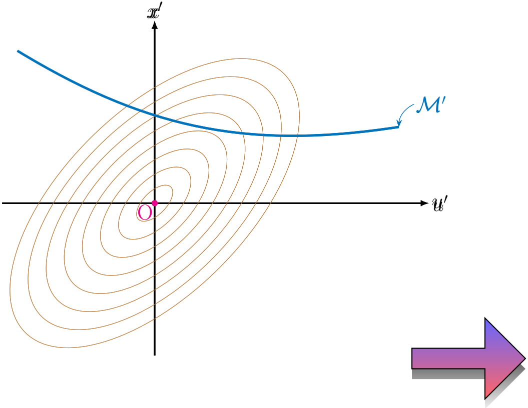


Linear Transformation W: $z' = W(z)$

Key Idea 2: Preconditioning...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz \rangle$$
$$\text{subject to} \quad z \in \mathcal{M}$$

$$\underset{z'}{\text{minimize}} \quad J'(z') = \frac{1}{2}\langle z', z' \rangle$$
$$\text{subject to} \quad z' \in \mathcal{M}'$$

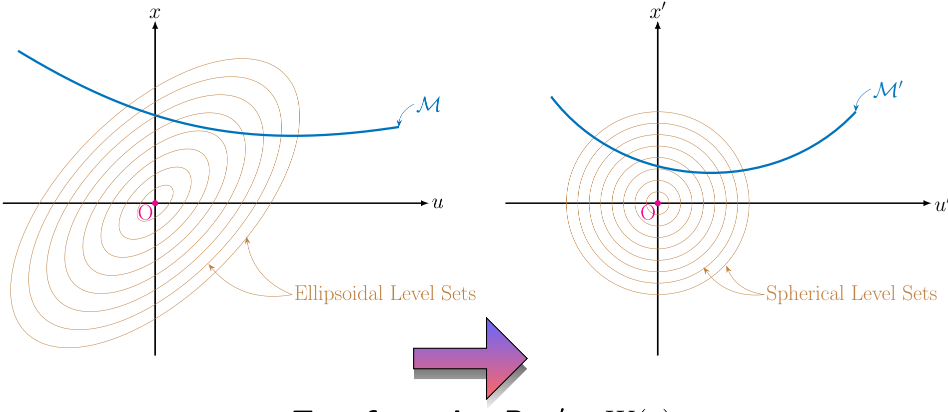


Transformation W: $z' = W(z)$

Key Idea 2: Preconditioning...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle$$
$$\text{subject to} \quad z \in \mathcal{M}$$

$$\underset{z'}{\text{minimize}} \quad J'(z') = \frac{1}{2}\langle z', z'\rangle$$
$$\text{subject to} \quad z' \in \mathcal{M}'$$



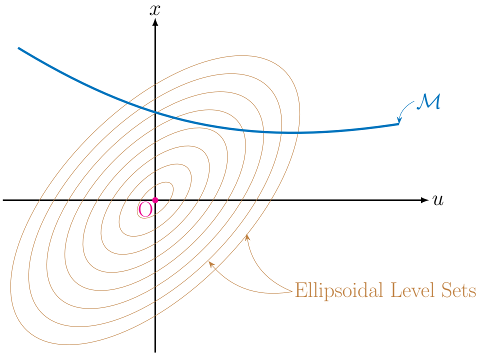
Transformation P: $z' = W(z)$

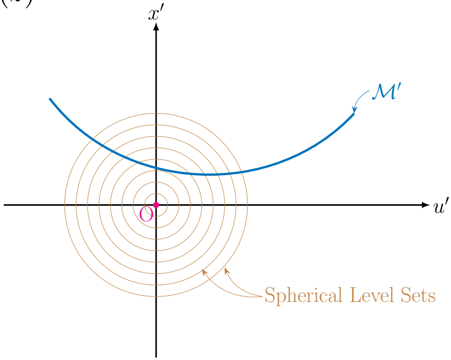
Key Idea 2: Preconditioning...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz \rangle$$
$$\text{subject to} \quad z \in \mathcal{M}$$

$$z' = W(z)$$

$$\underset{z'}{\text{minimize}} \quad J'(z') = \frac{1}{2}\langle z', z' \rangle$$
$$\text{subject to} \quad z' \in \mathcal{M}'$$





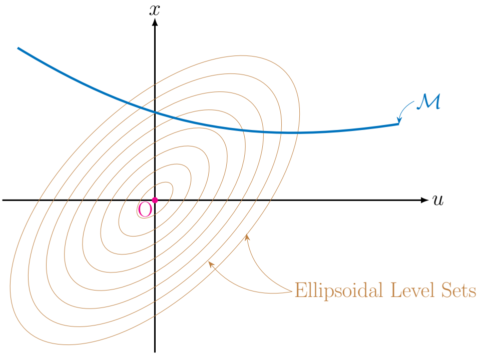
$$\begin{cases} \hat{z}'_{k+1} = z'_k - \alpha_k \Pi_{T_{z'_k}\mathcal{M}'}\left(\partial J'_k\right) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}} \circ W^{-1}\left(\hat{z}'_{k+1}\right) \end{cases}$$

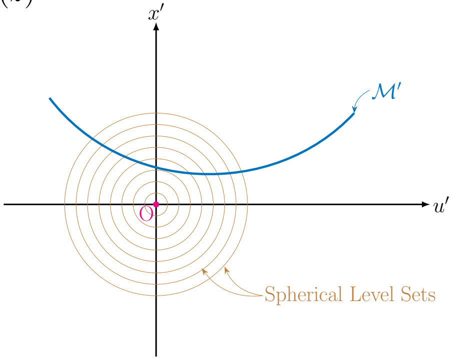
Key Idea 2: Preconditioning...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle$$
$$\text{subject to} \quad z \in \mathcal{M}$$

$$z' = W(z)$$

$$\underset{z'}{\text{minimize}} \quad J'(z') = \frac{1}{2}\langle z', z'\rangle$$
$$\text{subject to} \quad z' \in \mathcal{M}'$$



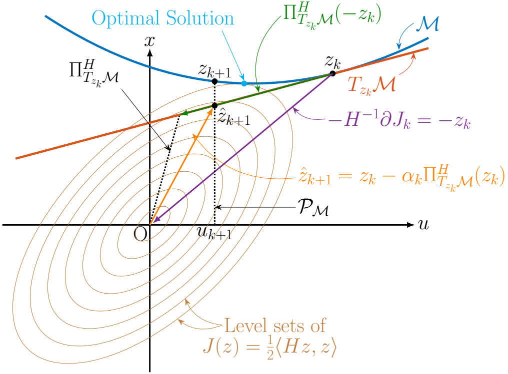


$$\begin{cases} \hat{z}_{k+1} = z_k - \alpha_k \Pi^{H}_{T_{z_k}\mathcal{M}}\left(H^{-1}\partial J_k\right) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1}) \end{cases}$$

$$\begin{cases} \hat{z}'_{k+1} = z'_k - \alpha_k \Pi_{T_{z'_k}\mathcal{M}'}\left(\partial J'_k\right) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}} \circ W^{-1}\left(\hat{z}'_{k+1}\right) \end{cases}$$

Key Idea 2: Preconditioning...

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle$$
$$\text{subject to} \quad z \in \mathcal{M}$$



$$\begin{cases} \hat{z}_{k+1} = z_k - \alpha_k \Pi^{H}_{T_{z_k}\mathcal{M}}\left(H^{-1}\partial J_k\right) \\ z_{k+1} = \mathcal{P}_\mathcal{M}(\hat{z}_{k+1}) \end{cases}$$

Computational Load

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz \rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$

$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

Computational Load

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$

$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

$$\text{PCGD}: \begin{cases} \hat{z}_{k+1} = z_k + \alpha_k \tilde{z}_k; \qquad \tilde{z}_k = -\Pi^{H}_{T_{z_k}\mathcal{M}}\left(H^{-1}\partial J_k\right) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1}) \end{cases}$$

Computational Load

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$

$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

$$\text{PCGD}: \begin{cases} \hat{z}_{k+1} = z_k + \alpha_k \tilde{z}_k; & \tilde{z}_k = -\Pi^H_{T_{z_k}\mathcal{M}}\left(H^{-1}\partial J_k\right) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1}) & \end{cases}$$

- $\Pi^H_{T_{z_k}\mathcal{M}}$: Solve a linear two point boundary value problem for $\tilde{z}_k := (\tilde{x}_k, \tilde{u}_k)$

$$\frac{d}{dt}\begin{bmatrix} \tilde{x}_k \\ \tilde{\lambda}_k \end{bmatrix} = \begin{bmatrix} A_k & -B_k R^{-1} B_k^* \\ -Q & -A_k^* \end{bmatrix}\begin{bmatrix} \tilde{x}_k \\ \tilde{\lambda}_k \end{bmatrix} - \begin{bmatrix} B_k u_k \\ Q x_k \end{bmatrix}; \quad \begin{bmatrix} \tilde{x}_k(0) \\ \tilde{\lambda}_k(T) \end{bmatrix} = 0$$

$$\tilde{u}_k = -\left(u_k + R^{-1} B_k^* \tilde{\lambda}_k\right)$$

where $A_k := \partial_x f(x_k, u_k)$ and $B_k := \partial_u f(x_k, u_k)$

Computational Load

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz \rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$

$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

$$\text{PCGD :} \begin{cases} \hat{z}_{k+1} = z_k + \alpha_k \tilde{z}_k; \qquad \tilde{z}_k = -\Pi^H_{T_{z_k}\mathcal{M}}\left(H^{-1}\partial J_k\right) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1}) \end{cases}$$

- $\Pi^H_{T_{z_k}\mathcal{M}}$: Solve a linear two point boundary value problem
- $\mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1})$: Solve the system dynamics

Computational Load

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz \rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$

$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

$$\text{PCGD}: \begin{cases} \hat{z}_{k+1} = z_k + \alpha_k \tilde{z}_k; & \tilde{z}_k = -\Pi^{H}_{T_{z_k}\mathcal{M}}\left(H^{-1}\partial J_k\right) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1}) \end{cases}$$

- $\Pi^{H}_{T_{z_k}\mathcal{M}}$: Solve a linear two point boundary value problem
- $\mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1})$: Solve the system dynamics
- No Costate Equation!

Computational Load

$$\underset{z}{\text{minimize}} \quad J(z) = \frac{1}{2}\langle z, Hz\rangle \qquad H := \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$

$$\text{subject to} \quad \dot{x}(t) = f\big(x(t), u(t)\big); \quad x(0) = x_0$$

$$\text{PCGD :} \quad \begin{cases} \hat{z}_{k+1} = z_k + \alpha_k \tilde{z}_k; & \tilde{z}_k = -\Pi^{H}_{T_{z_k}\mathcal{M}}\left(H^{-1}\partial J_k\right) \\ z_{k+1} = \mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1}) \end{cases}$$

- $\Pi^{H}_{T_{z_k}\mathcal{M}}$: Solve a linear two point boundary value problem
- $\mathcal{P}_{\mathcal{M}}(\hat{z}_{k+1})$: Solve the system dynamics
- No Costate Equation!
- No second derivatives of the dynamics!

Example: Comparison with the Standard Gradient Descent

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T \left[|\psi(t)^* Q \psi(t)| + Ru^2(t)\right] dt$$

$$\text{subject to} \quad i\hbar \frac{d}{dt}\psi(t) = [H_0 + Vu(t)]\psi(t); \quad \psi(0) = \psi_0$$

$$H_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1.7 \end{bmatrix}; \quad V = \begin{bmatrix} 0 & 1 & 0.32 \\ 1 & 0 & 0.95 \\ 0.32 & 0.95 & 0 \end{bmatrix}; \quad Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad R = \hbar = 1$$

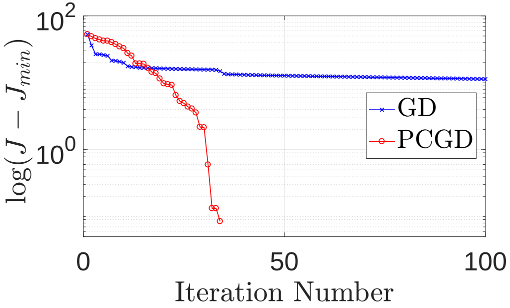
[2] Grivopoulos, Symeon. Optimal control of quantum systems. University of California, Santa Barbara, 2005.

Example: Comparison with the Standard Gradient Descent

$$\underset{x,u}{\text{minimize}} \quad J(x,u) = \frac{1}{2}\int_0^T \left[|\psi(t)^* Q \psi(t)| + Ru^2(t)\right] dt$$

$$\text{subject to} \quad i\hbar \frac{d}{dt}\psi(t) = [H_0 + Vu(t)]\psi(t); \quad \psi(0) = \psi_0$$

$$H_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1.7 \end{bmatrix}; \quad V = \begin{bmatrix} 0 & 1 & 0.32 \\ 1 & 0 & 0.95 \\ 0.32 & 0.95 & 0 \end{bmatrix}; \quad Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad R = \hbar = 1$$



[2]Grivopoulos, Symeon. Optimal control of quantum systems. University of California, Santa Barbara, 2005.

Concluding Remarks & Future Work

- PCGD can be shown to be a Quasi-Newton method
- PCGD inherits attractive properties of both first and second order methods:
  - Guaranteed to converge to a local minimum
  - Exhibits fast convergence rate near the optimum

Future work: How to include inequality constraints?

Concluding Remarks & Future Work

- PCGD can be shown to be a Quasi-Newton method
- PCGD inherits attractive properties of both first and second order methods:
  - Guaranteed to converge to a local minimum
  - Exhibits fast convergence rate near the optimum

Future work: How to include inequality constraints?

Thank you