Output-feedback Synthesis Orbit Geometry: Quotient Manifolds and LQG Direct Policy Optimization

Spencer Kraisler, Mehran Mesbahi

RAIN Lab
William E. Boeing Department of Aeronautics & Astronautics
University of Washington

December 2024

What is Direct Policy Optimization (PO)?

- **Goal:** bridge control theory and RL
- Design controllers via policy gradient methods
  - **Novel focus:** optimizer performance $\Longrightarrow$ study geometry of policy space + performance measure
- **Idea:** if training in real time, we need to know when policy will be safe

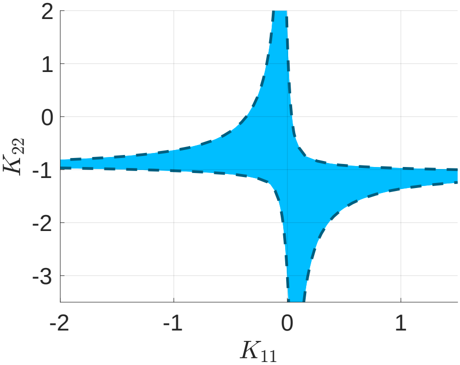


Figure: The set of stabilizing 2×2 **diagonal** feedback matrices K (Talebi, 2024)

Previous Results

LQR

- **Domain:** stabilizing static feedback matrices
- analytic, non-convex, gradient dominant
- global convergence under gradient descent (GD) with linear rate

LQG

- **Domain:** stabilizing **dynamic** linear controllers
- analytic, non-convex, non-strict saddle points, degenerate stationary points
- Sublinear convergence rate under GD
- No local convergence guarantee

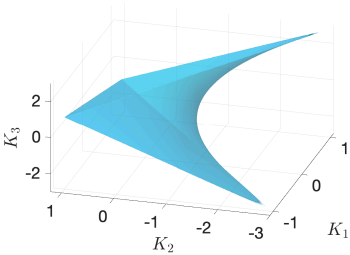


Figure: Set of stabilizing static controllers

Questions: Why sub-linear convergence rate? Why no convergence guarantee?

Stochastic LTI Systems

Consider the stochastic linear system

$$\dot{x} = Ax + Bu + w,$$
$$y = Cx + v$$

in feedback with a dynamic linear controller

$$\dot{\xi} = A_{\mathsf{K}}\xi + B_{\mathsf{K}}y,$$
$$u = C_{\mathsf{K}}\xi$$
$$\mathsf{K} := (A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}})$$

Controller space: $\widetilde{\mathcal{C}}_n \underset{\text{open}}{\subset} \mathbb{R}^{n^2+nm+np}$, stabilizing full-order minimal (i.e. controllable + observable) controllers

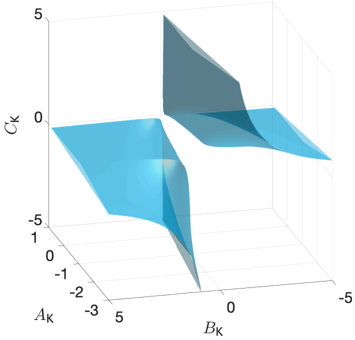


Figure: Illustration of the set of dynamic stabilizing policies $\widetilde{\mathcal{C}}_1$ for an LTI system with $A = 1.1$ and $B = C = 1$, resulting in two path-connected components

Stochastic LTI System (Cont.)

Important Coordinate-transformation: $S\xi = \eta$, $\mathcal{T}_S(\mathsf{K}) = (SA_\mathsf{K}S^{-1}, SB_\mathsf{K}, C_\mathsf{K}S^{-1})$

$$\begin{cases} \dot{\xi} = A_\mathsf{K}\xi + B_\mathsf{K}y, \\ u = C_\mathsf{K}\xi \end{cases} \implies \begin{cases} \dot{\eta} = SA_\mathsf{K}S^{-1}\eta + SB_\mathsf{K}y, \\ u = C_\mathsf{K}S^{-1}\eta \end{cases}$$

LQG Cost:

$$\widetilde{J}(\mathsf{K}) := \lim_{T\to\infty} \mathbb{E}_w \frac{1}{T}\int_0^T (x^\mathsf{T}Qx + u^\mathsf{T}Ru)dt$$

Theorem (Zheng, Tang, & Li, 2021)

$\widetilde{J} : \widetilde{\mathcal{C}}_n \to \mathbb{R}$ is analytic, non-convex, all minima are global, admits saddle points. Also, coordinate-invariance: $\widetilde{J}(\mathcal{T}_S(\mathsf{K})) = \widetilde{J}(\mathsf{K})$.

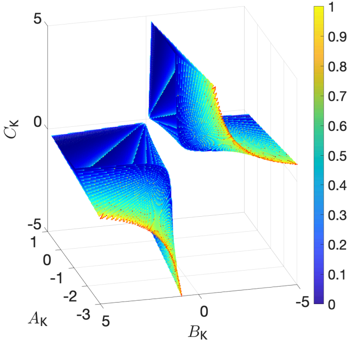


Figure: A colored plot of $\widetilde{J}(\cdot)$ over $\widetilde{\mathcal{C}}_1$.

Direct PO Re-visited

Coordinate-invariance $\implies$ n^2 dimensions of REDUNDENCY

An orbit: $[\mathrm{K}] := \{\mathcal{T}_S(\mathrm{K}) : S \in \mathrm{GL}(n)\}$

Problem

Minimize $\widetilde{J}$ over $\widetilde{\mathcal{C}}_n$ in a fast way, with at least linear rate with local convergence guarantee that takes advantage of coordinate-invariance property

Solution: Reformulate $\widetilde{\mathcal{C}}_n$ as a Riemannian manifold such that "$\nabla \widetilde{J}(\mathrm{K}) \perp [\mathrm{K}]$"

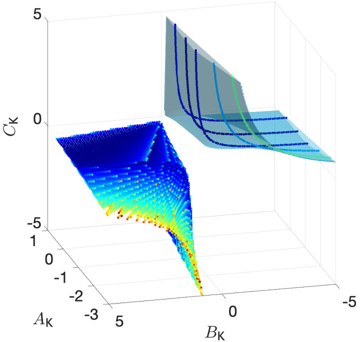


Figure: A colored plot of $\widetilde{J}(\cdot)$ over $\widetilde{\mathcal{C}}_1$ but with only a few orbits colored in one connected component.

Smooth Manifold

Smooth manifold

- A space which is *locally Euclidean*
  - $\widetilde{\mathcal{C}}_n$, Unit quaternions $\mathcal{Q}$
- Compatible with calculus
- Tangent space $T_x\mathcal{M} \cong \mathbb{R}^n$
  - $T_{\mathsf{K}}\widetilde{\mathcal{C}}_n \cong \mathbb{R}^{n^2+nm+np}$,
    $T_q\mathcal{Q} = \{v \in \mathbb{R}^4 : v^\mathsf{T} q = 0\}$

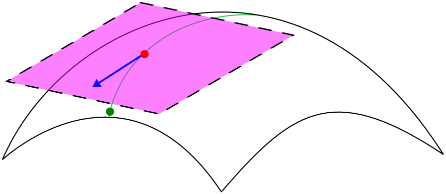


Figure: Every point admits a vector space of *tangent vectors*

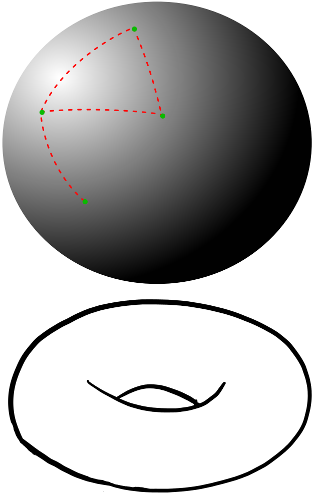


Figure: Smooth manifold examples

Example:

$$f(x) := x^4, \quad \langle v, w\rangle_x := v \cdot x^2 \cdot w$$
$$\overline{\nabla} f(x) = 4x^3, \quad \nabla f(x) = 4x$$

Riemannian metric:

- Inner product $\langle \cdot, \cdot \rangle_x$ on each tangent space $T_x\mathcal{M}$
- length, angle, area, gradient, Hessian

Riemannian gradient:
$G(x)\nabla f(x) = \overline{\nabla} f(x)$, where $\overline{\nabla} f$ is the ordinary (Euclidean) gradient

Intuition: Riemannian gradient is a pre-conditioning on the gradient field $G(x)^{-1}\overline{\nabla} f(x)$ (example: barrier method)

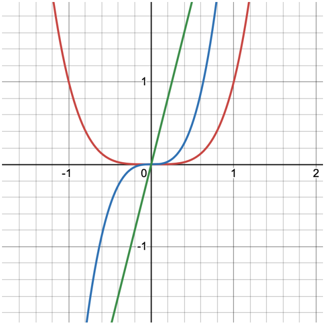


Figure: The graphs of $f(x)$, $\overline{\nabla} f(x)$, and $\nabla f(x)$.

Retraction: $R_x : T_x\mathcal{M} \to \mathcal{M}$

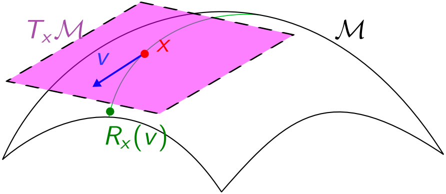


Figure: Retraction visualization

Riemannian Gradient Descent (RGD):

$$x_{k+1} = R_{x_k}(-\alpha \nabla f(x_k))$$

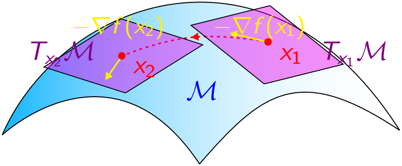


Figure: RGD visualization; here, $x_2 = R_{x_1}(-\alpha \nabla f(x_1))$.

Important: The right Riemannian metric can speed up convergence rate of RGD from sub-linear to linear!

The Krishnaprasad-Martin Metric

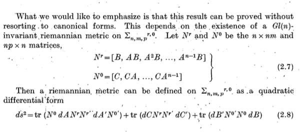
Intuition for the *right* Riemannian metric $\langle \mathbf{V}, \mathbf{W} \rangle_{\mathsf{K}} : T\widetilde{\mathcal{C}}_n \times T\widetilde{\mathcal{C}}_n \to \mathbb{R}$:

1. Should be coordinate-invariant: $\langle \mathcal{T}_S(\mathbf{V}), \mathcal{T}_S(\mathbf{W}) \rangle_{\mathcal{T}_S(\mathsf{K})} = \langle \mathbf{V}, \mathbf{W} \rangle_{\mathsf{K}}$
2. $\langle \mathbf{V}, \mathbf{W} \rangle_{\mathsf{K}}$ should explode as K becomes less stablizing (mimicking barrier method techniques)

Definition (Krishnaprasad-Martin metric)

$$
\begin{aligned}
\langle \mathbf{V}, \mathbf{W} \rangle_{\mathsf{K}}^{\mathrm{KM}} := {} & c_1 \operatorname{tr}\left(W_o(\mathsf{K}) E(\mathbf{V}) W_c(\mathsf{K}) E(\mathbf{W})^T\right) \\
& + c_2 \operatorname{tr}\left(F(\mathbf{V})^T W_o(\mathsf{K}) F(\mathbf{W})\right) \\
& + c_3 \operatorname{tr}\left(G(\mathbf{V}) W_c(\mathsf{K}) G(\mathbf{W})^T\right)
\end{aligned}
$$

where $c_1 > 0$, $c_2, c_3 \geq 0$, and $W_c(\cdot), W_o(\cdot)$ are closed-loop controllability/observability Grammians



What we would like to emphasize is that this result can be proved without resorting to canonical forms. This depends on the existence of a $Gl(n)$-invariant riemannian metric on $\Sigma_{n,m,p}{}^{r,0}$. Let N^r and N^0 be the $n \times nm$ and $np \times n$ matrices,

$$
\left.\begin{aligned}
N^r &= [B, AB, A^2B, \ldots, A^{n-1}B] \\
N^0 &= [C, CA, \ldots, CA^{n-1}]
\end{aligned}\right\} \tag{2.7}
$$

Then a riemannian metric can be defined on $\Sigma_{n,m,p}{}^{r,0}$ as a quadratic differential form

$$
ds^2 = \operatorname{tr}(N^0\, dA\, N^r N^{r\prime}\, dA' N^{0\prime}) + \operatorname{tr}(dC\, N^r N^{r\prime}\, dC') + \operatorname{tr}(dB' N^{0\prime} N^0\, dB) \tag{2.8}
$$

Figure: Developed in 1983 to study manifold of stable LTI systems, involved control theorists such as Kalman, Tannenbaum, and Brockett

The Algorithm and Convergence Analysis

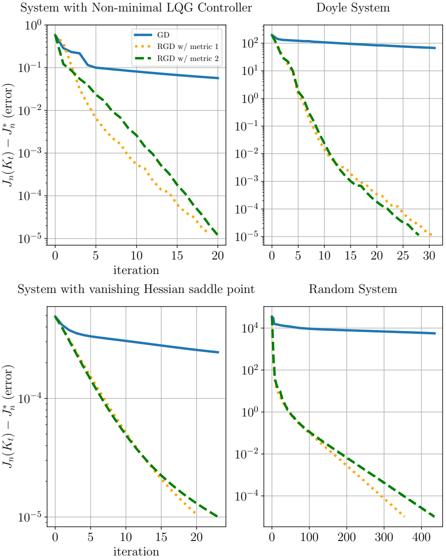
RGD of $(\widetilde{\mathcal{C}}_n, \langle\cdot,\cdot\rangle^{\mathrm{KM}}, +)$ over $\widetilde{J}(\cdot)$ with fixed step size:

$$\mathsf{K}_{t+1} = \mathsf{K}_t - \alpha \nabla \widetilde{J}(\mathsf{K}_t) \qquad (3)$$

Theorem (Kraisler and Mesbahi, 2024)

Suppose the LQG controller is controllable + observable, and $\mathrm{null}\nabla^2 \widetilde{J}(\mathsf{K}^*) = T_{\mathsf{K}^*}[\mathsf{K}^*]$. *Then there exists* $\alpha > 0$ *and a neighborhood* U *of* K^* *such that the sequence defined by (3) with* $\mathsf{K}_0 \in U$ *exists and converges to* $[\mathsf{K}^*]$ *with at least linear rate.*

Summary: theoretical gauarantee on local convergence + linear rate



An Interpretation: *Smooth Quotient Manifolds*

- Orbit: $[\mathsf{K}] = \{\mathcal{T}_S(\mathsf{K}) : S \in \mathrm{GL}_n\}$
- smooth quotient manifold:
  $\mathcal{C}_n := \{[\mathsf{K}] : \mathsf{K} \in \widetilde{\mathcal{C}}_n\}$
- $\dim(\widetilde{\mathcal{C}}_n) = n^2 + nm + np$ and
  $\dim(\mathcal{C}_n) = nm + np$
- **Important:** RGD over $\widetilde{\mathcal{C}}_n \iff$ RGD over the (smaller dimensional) $\mathcal{C}_n$

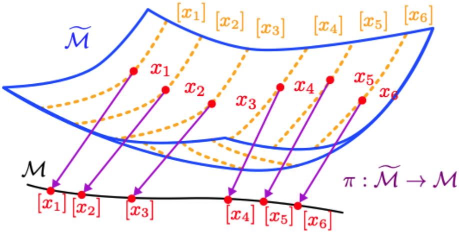


Figure: Visualization of a quotient manifold

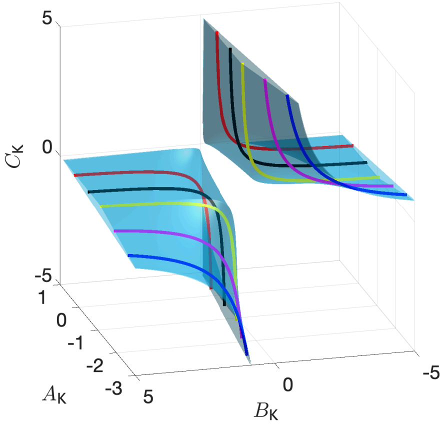


Figure: The orbits of $\widetilde{\mathcal{C}}_1$ colored

Conclusion

Summary:

- RGD order of magnitude faster than GD
- Local convergence guarantee
- Interpretation: RGD over the much smaller quotient controller manifold

Future directions:

- 2nd-order methods
- Studing the topology and geometry of $\mathcal{C}_n$
- $\mathcal{H}_\infty$ non-smooth optimization over the controller quotient manifold $\mathcal{C}_n$
- Data-driven synthesis of filters (Kalman filter)