Università degli Studi di Verona

Master's Degree - Mathematics
Course Report

Data-Driven Methods for Optimal Control

Student:
Michele Benedetti
Student ID:
VR488172

Supervisor(s):
Prof. Dante Kalise
Prof. Giacomo Albi

December 11, 2022

Contents

1 Control Fundamentals

General Control Theory has its roots in the study of dynamical systems. Consider indeed the following system:

$$\begin{cases} \dot{x} = f(x) \\ x(0) = x_0 \end{cases} \tag{1}$$

where $x : \mathbb{R} \to \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$.

Assume to introduce a generic **control variable** u in such a way that the system can be then written as follows:

$$\begin{cases} \dot{x} = f(x, u) \\ x(0) = x_0 \end{cases} \tag{2}$$

where $x : \mathbb{R} \to \mathbb{R}^n$, $u : \mathbb{R} \to \mathbb{R}^n$ and $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$.

Our main target will be the study of the **synthesis** of u. Let, for example, x be a solution of the of the dynamical system (1) such that it induces oscillatory trajectories. We might want to find a suitable control u that, when inserted into (2), renders the trajectories steady:

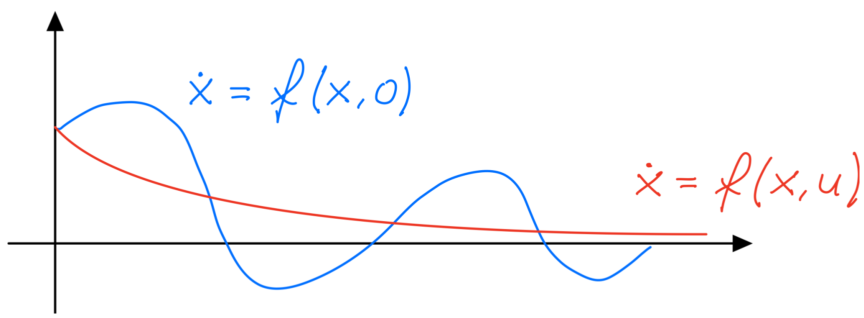

Figure 1: Steadiness

1.1 Example: Motorized Pendulum

Consider a generic pendulum system (m, θ) coupled with a motor that is able to generate an angular displacement via a force u:

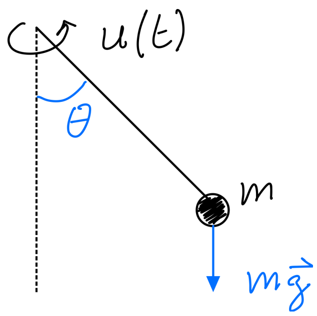

Figure 2: Pendulum

From basic Newtonian Mechanics we have the following non-linear ODE system for the dynamics:

$$\begin{cases} m \cdot \ddot{\theta}(t) + m \cdot g \cdot \sin(\theta(t)) = u(t) \\ \theta(0) = \theta_0 \end{cases}$$

We can re-write the system in a static-space form as follows: first, we introduce the **state variables**

$$\begin{cases} x_1(t) = \theta(t) \\ x_2(t) = \dot{\theta}(t) \end{cases}$$

assuming now that $m = g = 1$, we obtain:

$$\dot{X} = \frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -\sin(x_1) + u(t) \end{pmatrix}$$

where u is the **control variable**.

This non-linear system has 2 equilibria: $\theta = 0$ (stable) and $\theta = \pi$ (unstable). This means that, whatever the initial condition, for a sufficiently large t we'll always have that the system converges to the stable equilibrium $\theta = 0$. Assume however that we want to force the system to stabilize at the unstable equilibrium $\theta = \pi$ using the force u. We linearize the system taking the following approximations:

$$\theta \approx \pi \implies \sin(\theta) \approx -(\theta - \pi)$$

Therefore we introduce $\phi := \theta - \pi$ in order to get

$$\ddot{\phi}(t) - \phi(t) = u(t)$$

also called the **feedback control**.

We want to stabilize the system at the unstable equilibrium, so we take the **control law** to be defined as $u(t) = -\alpha \cdot \phi(t)$, where $\alpha > 0$. Now we write the **closed loop system** as:

$$\ddot{\phi} - \phi = -\alpha \cdot \phi$$

We can now use linear stability analysis to study the effect of the control law. A very simple improvement of the above control law is given by the possible introduction of a **damping term** by taking

$$u = -\alpha \cdot \phi - \beta \cdot \dot{\phi}$$

where $\beta > 0$ is the **damping**. Finally, we can also introduce an integral term to obtain a so called **PID** (**Proportional - Integral - Derivative**):

$$u = -\alpha \cdot \phi - \beta \cdot \dot{\phi} - \gamma \cdot \int_0^T \phi(t)dt$$

2 Introduction to Optimal Control

Optimal Control (also referred to as Dynamic Optimization) problems can be formulated as follows:

$$\min_u \int_0^T L(x,u)dt + V(x(T)) \tag{3}$$

$$\text{s.t.} \quad \begin{cases} \dot{x} = f(x,u), & x \in \mathbb{R}^n \\ x(0) = x_0, & x_0 \in \mathbb{R}^n \end{cases}$$

where $L(\cdot,\cdot)$ is called the **optimization term** and $V(\cdot)$ is called the **penalty term**.

A solution u^* of (3) is called an **optimal control** and the trajectory x^* induced by u^* is called an **optimal trajectory**. Note that we can discretize the problem as follows:

$$\begin{aligned} \dot{x} = f(x,u) &\implies x^{k+1} = x^k + \Delta t \cdot f(x^k, u^k) \\ &\implies x(t) \to \{x^k\}, \quad u(t) \to \{u^k\} \\ &\implies \int_0^T L(x,u)dt \approx \sum_{i=1}^{N_T} L(x^k, u^k) \end{aligned}$$

2.1 Some types of Control

We now briefly present 3 possible types of controls:

- ∞-Horizon Optimal Control:

  We take $T = +\infty$, $V \equiv 0$, so we obtain the following formulation for (3):

  $$\min_u \int_0^T L(x,u)dt \quad \text{s.t.} \quad \dot{x} = f(x,u)$$

  $$\implies \exists (x^*, u^*) \quad \text{s.t.} \quad L(x^*, u^*) \to_{t\to+\infty} 0$$

  We could for example take $L(x,u) := ||x||_2^2 + ||u||_2^2$

- Linear-Quadratic Control:

  We take $L(x,u) := x^T Q_x x + u^T Q_u u$ (**quadratic cost**), $V(x(T)) = x^T(T)P_1x(T)$ so we obtain the following formulation for (3):

  $$\min_u \int_0^{+\infty} (x^T Q_x x + u^T Q_u u) \cdot dt + x^T(T)P_1x(T) \quad \text{s.t.} \quad \begin{cases} \dot{x} = Ax + Bu, \\ Q_x \geq 0, \quad Q_u \geq 0 \end{cases}$$

  $$\implies L(x,u) \to_{t\to+\infty} 0$$

- Time-optimal Control:

  We take $L = 1$, $V \equiv 0$ and leave T free, we obtain the following formulation for (3):

$$\min_{(u,T)} T = \int_0^T 1 \cdot dt \quad \text{s.t.} \quad \begin{cases} \dot{x} = f(x,u), \\ x(0) = x_0, \\ x(T) = x_d, \\ u \in \mathcal{U} \end{cases}$$

  where x_d is the **desired final state** and $\mathcal{U}$ is compact. There are 2 main difficulties with this formulation:

  - it is an optimization problem w.r.t. u and T
  - we don't know if $\exists T$ such that $x(T) = x_d$

  Optimality conditions are given by **Pontryagin's Maximum Principle**: given the optimal control problem

$$\min_u \int_0^T L(x,u)dt + V(x(T)) \quad \text{s.t.} \quad \begin{cases} \dot{x} = f(x,u), \quad x \in \mathbb{R}^n \\ x(0) = x_0, \quad x_0 \in \mathbb{R}^n \end{cases}$$

  we construct the **Hamiltonian** of the system as follows:

$$\mathcal{H}(x,u,\lambda) = \mathcal{L}(x,u) + \lambda^T \cdot f(x,u)$$

  where λ is the **adjoint variable** of $\mathcal{H}$. The stationary conditions for $\mathcal{H}$ are:

$$\nabla_x(H) = \nabla_u(H) = \nabla_\lambda(H) = 0$$

  By Pontryagin's Maximum Principle (**P.M.P.**) there $\exists f$ s.t. (x^*, u^*) is optimal i.e. there $\exists \lambda^* \in \mathbb{R}^n$, $\gamma^* \in \mathbb{R}^q$ s.t.:

$$\begin{cases} \dot{x} = \partial_\lambda \mathcal{H}, \\ -\dot{\lambda}_i = \partial_{x_i} \mathcal{H} \quad \text{(adjoint equation)}, \\ x(0) = x_0, \\ \psi(x(T)) = 0, \\ \lambda(T) = \partial_x V(x(T)) + \gamma^T \partial_x \psi, \\ \mathcal{H}(x^*, u^*, \lambda^*) \leq \mathcal{H}(x^*, u, \lambda^*) \quad \forall u, \forall t \quad (\Longleftrightarrow u^* = \arg\min_u \mathcal{H}(x^*, u, \lambda^*)) \end{cases}$$

2.2 Example: 2 point BPV (TPBPV)

Consider the following problem:

$$\begin{cases} \dot{x} = a \cdot x + b \cdot u \\ x_0 = x(t_0) \end{cases} \tag{4}$$

We have:

$$J := \frac{1}{2} \int_0^T u^2 dt + \frac{1}{2} c \cdot x(T)^2, \quad c > 0$$

Therefore:

$$\implies L = \frac{1}{2} u^2, \quad V(x) = \frac{1}{2} c \cdot x(T)^2$$

Which implies the following expression for the Hamiltonian of the system:

$$\implies \mathcal{H}(x, u, \lambda) = \frac{1}{2} \cdot u^2 + \lambda \cdot (a \cdot x + b \cdot u)$$

Applying **P.M.P.** conditions to this Hamiltonian we obtain the following problem:

$$\begin{cases} \dot{x} = a \cdot x + b \cdot u \\ x_0 = x(t_0) \\ -\dot{\lambda} = \lambda \cdot a \\ \lambda(T) = c \cdot x(T) \end{cases} \tag{5}$$

And therefore:

$$\begin{aligned} &\implies \mathcal{H}(x^*, u^*, \lambda^*) \leq \mathcal{H}(x^*, u, \lambda^*) \quad \forall t \in [0, T] \\ &\iff u^* = \arg\min_u \mathcal{H}(x^*, u, \lambda^*) = \arg\min_u \frac{1}{2} u^2 + \lambda^* \cdot a \cdot x^* + \lambda^* \cdot b \cdot u \\ &\implies u^* = -\lambda \cdot b \end{aligned}$$

Plugging this expression for u^* into (5) leads to the 2 point Boundary Value Problem:

$$\begin{cases} \dot{x} = a \cdot x - b^2 \cdot \lambda \\ x_0 = x(t_0) \\ -\dot{\lambda} = \lambda \cdot a \\ \lambda(T) = c \cdot x(T) \end{cases} \tag{6}$$

Which consists in an equation **forwards** in time for x and an equation **backwards** in time for λ. The system can be represented in the following way:

$$\begin{pmatrix} x(0) \\ ? \end{pmatrix} \underset{-\dot{\lambda} = a\lambda}{\overset{\dot{x} = ax - b^2\lambda}{\rightleftarrows}} \begin{pmatrix} ? \\ \lambda(T) \end{pmatrix}$$

In this particular case we can integrate and obtain:

$$\lambda(t) = c \cdot x(T) \cdot e^{a(T-t)} \implies \dot{x} = a \cdot x - b^2 c \cdot x(T) \cdot e^{a(T-t)} \quad \ldots$$

2.3 Reduced Gradient

Given an optimal control problem, numerically we have 3 alternatives:

- Proceed as above and obtain a TPBPV. However this procedure **is not always possible and depends on** u**!**
- Discretize everything and obtain the following:

$$\begin{cases} x^{k+1} = x^k \Delta t \cdot f(x^k, u^k), & x^0 = ... \\ \lambda^{k+1} = \lambda^k - \Delta t \cdot ..., & \lambda^N = ... \end{cases} \implies F(\vec{x}, \vec{\lambda}) = 0$$

  which can be resolved using Newton's Method...

- **Reduced Gradient**: we have the following

$$\min_u J(x,u) \; (= \int_0^T L(x,u)dt + V(x(T)) \;) \quad \text{s.t.} \quad \begin{cases} \dot{x} = f(x,u), & x \in \mathbb{R}^n \\ x(0) = x_0, & x_0 \in \mathbb{R}^n \end{cases} \tag{7}$$

  if we fix x_0 we have the so-called **control-to-state map**:

$$u(\cdot) \longmapsto x(t) := x(u)$$

  so in this case the problem (7) becomes:

$$\min_u J(x(u), u) \quad \text{s.t.} \quad \begin{cases} \dot{x} = f(x,u) \\ x(0) = x_0 \end{cases} \tag{8}$$

  given an initial guess u^0 of $u(t) \; \forall t \in [0,T]$, we can use a gradient method:

$$u^0 \rightarrow_{k \to +\infty} u^* \implies u^{k+1} = u^k - \delta^k \cdot \nabla_u J(u^k)$$

  Using Calculus of Variations, we can prove that we actually have the following:

$$\nabla_u J = \nabla_u \mathcal{H}(x, u, \lambda)$$

  To compute $\nabla_u J(u^k)$ we can use the following algorithm:

  1. u^k given
  2. $\dot{x} = f(x,u) \longrightarrow$ we obtain x^k integrating **forwards**
  3. $-\dot{\lambda} = \partial_x \mathcal{H} \longrightarrow$ we obtain λ^k integrating **backwards**
  4. $\nabla_u J(u^k) = \nabla_u \mathcal{H}(x^k, u^k, \lambda^k)$

  u is still continuous at this level, so we nave to discretize it further: $u^k \longrightarrow \{u_i^k\}$ ... Moreover, note that in this case we don't have any constraints on u, but that is not always the case!

2.4 Example: Control-Affine Setting

Consider the following problem:

$$\dot{x} = f(x) + g(x) \cdot u$$

In this setting P.M.P. is a necessary condition, so we have:

$$\min_u J(x(u), u) \longrightarrow u^k(t) \in [0, T] \text{ which is the optimal control signal}$$

This map is uniquely determined by the initial condition $x(0)$:

$$x(0) \longrightarrow \min_u J(x(u), u) \longrightarrow u^k(t) \in [0, T]$$

However, in reality this almost never happens because we need to take errors/disturbances into account!

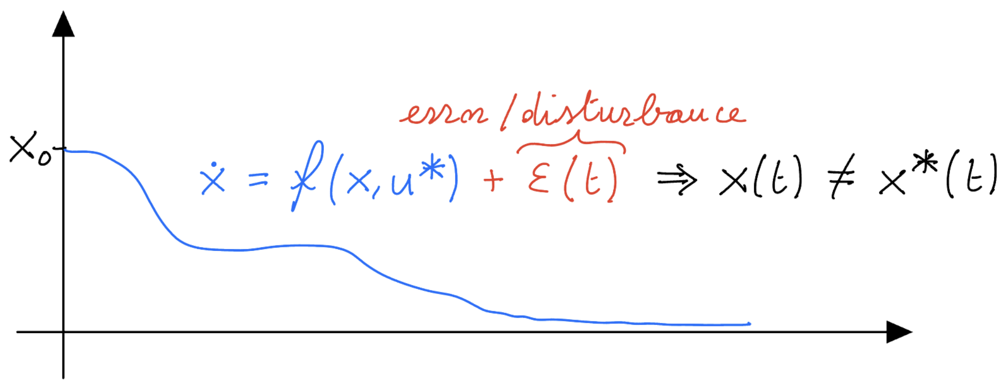

Figure 3: Errors and disturbances

We have the following distinction:

- Open-Loop Control:
$$u(t) \longrightarrow \dot{x} = f(x, u)$$
- Closed-Loop Control:
$$\dot{x} = f(x, u) \rightleftarrows u = F(x)$$
This is also referred to as model predictive control (**MPC**), $u = F(X)$ is called the **Feedback**.

3 Optimal Feedback

Consider the problem:

$$\begin{cases} \dot{x} = u(t), \ u \in \Omega = [-1, 1] \\ x(0) = x_0 \in \Omega = [-1, 1] \end{cases} \implies \partial\Omega = \{-1, 1\}$$

We want to minime the following:

$$\min T \quad \text{s.t} \quad x(T) \in \partial\Omega$$

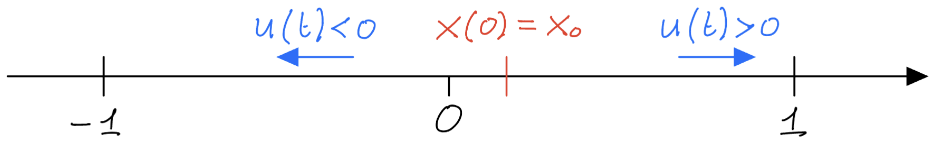

Figure 4: Problem dynamics

We have:

$$\min \int_0^T 1dt \implies \mathcal{H} := 1 + \lambda^T(u), \begin{cases} \dot{x} = u \\ -\dot{\lambda} = 0 \end{cases}$$

$$\implies u^* = \arg \min_{u\in[-1,1]} 1 + \lambda \cdot u = -\text{sgn}(\lambda) = \begin{cases} 1 & \lambda > 0 \\ -1 & \lambda < 0 \end{cases}$$

Which means that we have an **optimal solution in feedback form**:

$$u^*(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$

We now introduce the **arrival time** to $\partial\Omega$, $T(x)$:

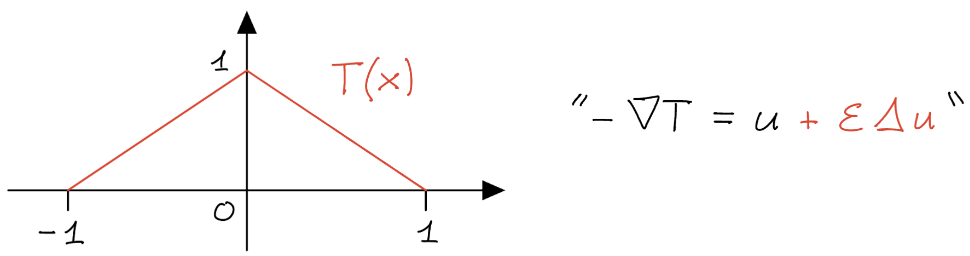

Figure 5: Arrival time

$T(x)$ solves the Hamilton-Jacobi-Bellman PDE (also called the Eikonal equation):

$$\begin{cases} ||\nabla T|| = 1 \\ T(-1) = T(1) = 0 \end{cases} \implies \begin{cases} T(x) := \inf_u J(\tilde{x}(u), u) \\ \tilde{x}(0) = x \end{cases}$$

Summarizing, we have the following scheme:

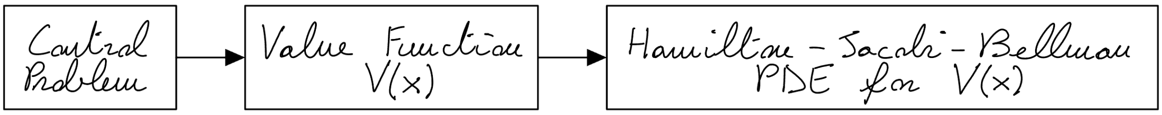

Figure 6: Optimal Feedback scheme

Provided that we can solve the PDE, we obtain the **optimal Feedback map**:

$$u^*(x) := -\nabla V(x)$$

If we have:

$$T(x) = \inf_u J(\tilde{x}(u), u)$$

then:

$$T(x) = \inf_u \{\tau + T(x(u,\tau))\} \; \forall x \in \Omega, \; \forall \tau \in [0, T(x)]$$

where $T(x(u,\tau))$ is **departing at** τ. Dividing by τ and taking $\tau \to 0$ we obtain:

$$-1 = \inf_u \{\frac{T(x(u,\tau)) - T(x)}{\tau}\} \longrightarrow_{\tau \to 0} \inf_u \{\nabla T \cdot f\} = \inf_u \{\nabla T \cdot \dot{x}\}$$
$$\iff 1 = \sup_u \{-\nabla T \cdot f\} = \sup_u \{-\nabla T \cdot \dot{x}\} \implies \begin{cases} \dot{x} = u, \; u \in [-1, 1] \\ 1 = ||\nabla u|| \end{cases}$$

So, if we solve the Hamilton-Jacobi-Bellman PDE for $T(x)$, we have the following:

$$u^*(x(t)) = \arg\max_u \{-\nabla T(x) \cdot f(x,u)\}$$

where $u^*(x(t))$ is the **current state**.

3.1 Link between P.M.P. and the H.-J.-B. PDE

Consider the following problem (**Finite Horizion/Unconstrained Problem**):

$$\min_u \int_t^T L(x(s), u(s))ds \quad \text{where} \quad L(x,u) := \ell(x) + \beta(u)^2 \tag{9}$$

$$\text{s.t.} \quad \begin{cases} \dot{x} = f(x) + g(x) \cdot u(s) \\ x(t) = x_0 \\ s \in [t, T] \end{cases}$$

We can study it in different ways:

1. **P.M.P.**
   Following the same procedure of section (2.1) we obtain the following:

$$\begin{cases} \mathcal{H} = L(x,u) + \lambda^T \cdot (f(x) + g(x) \cdot u) \\ \dot{x} = f(x) + g(x) \cdot u \\ x(t) = x_0 \end{cases}$$

   Which then implies

$$\implies \begin{cases} -\dot{\lambda} = \partial_x \mathcal{H} \\ \lambda(T) = 0 \end{cases} \implies u^* = \arg\min_u \{L(x,u) + \lambda^T \cdot (f(x) + g(x) \cdot u(s))\}$$

   Which means that finally we have the following:

$$\implies (u^*, x^*, \lambda^*(s)) \longrightarrow (u^*(t), x^*(t), \lambda^*(t)), \quad s \in [t, T]$$

2. **Hamilton-Jacobi-Bellman PDE**
   We follow the same steps of section (3) and obtain:

$$V(x,t) = \inf J(x,u) := \int_t^T L(x,u) ds = \int_t^T \ell(x) + \beta(u)^2 ds$$

   The Hamilton-Jacobi-Bellman PDE is:

$$\begin{cases} \partial_t V(x,t) = \frac{1}{4\beta} \nabla V^T \cdot g(x) \cdot g(x)^T \cdot \nabla V - \nabla V \cdot f(x) - \ell(x) \\ V(t,x) = 0 \end{cases}$$

   Which implies:

$$u^*(x,t) := \arg\min_u \ldots = -\frac{1}{2\beta} g(x)^T \cdot \nabla V(x,t) \quad \forall x \in \mathbb{R}^n, \forall t \in [0,T]$$

We want to study the existing link between the 2 approaches. P.M.P. conditions are far easier to compute numerically, but they are also completely dependent on the current state of the system. The Hamilton-Jacobi-Bellman PDE approach is extremely precise, but it involves solving a high-dimensional non-linear PDE! In this particular setting we have:

$$L(x,u) = \ell(x) + \beta(u)^2 \implies \dot{x} = f(x) + g(x) \cdot u \quad \text{(Control-Affine dynamics)}$$

Therefore, if we solve the P.M.P. conditions with initial condition $x(t_0) = \hat{x}$ and we obtain the **optimal trajectory** $(u^*(t), x^*(t), \lambda^*(t))$, we have:

$$\begin{cases} V(x^*(t), t) = \int_t^T \ell(x^*(s)) + \beta \cdot ||u^*(s)||^2 ds \\ \nabla V(x^*(t), t) = \lambda^*(t) \quad \forall t \in [t_0, T] \end{cases}$$

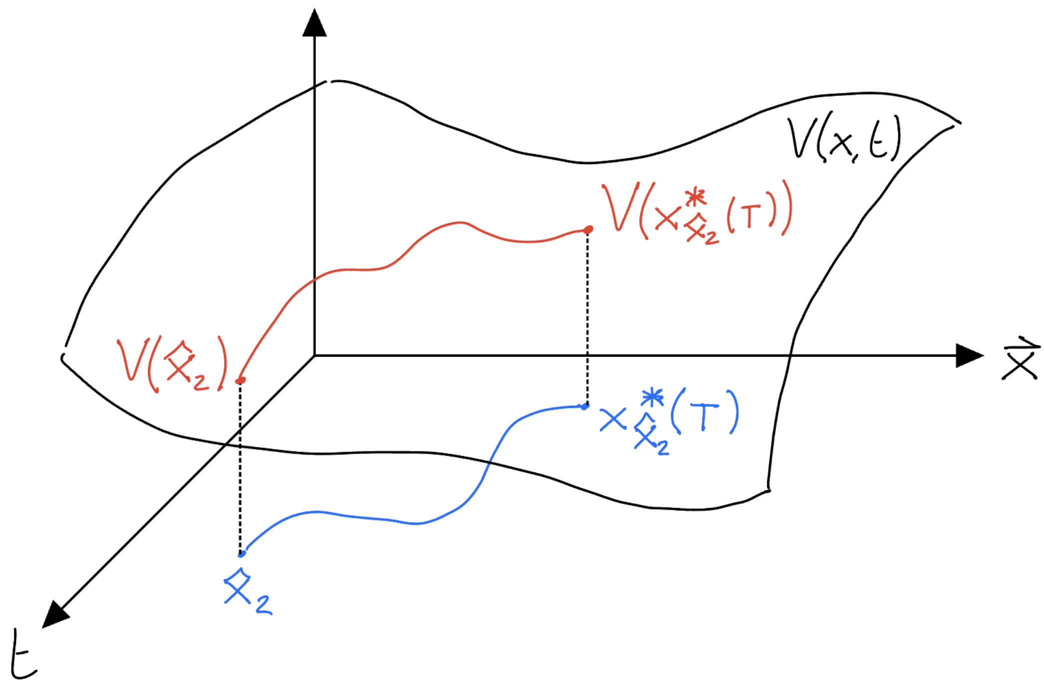

Figure 7: Control-Affine dynamics

Assuming $t = 0$ and approximating $V(x, 0)$ (i.e. solving the H.-J.-B. PDE) we obtain a **sub-optimal feedback**:

$$\tilde{u}(x) = -\frac{1}{2\beta} g(x)^T \cdot \nabla V(x, 0)$$

Finally, we can use the following algorithm to solve the problem numerically:

1. Generate samples $\{\hat{x}_i\}_{i=1}^{N_s}$ on $\mathbb{R}^n$
2. Solve the P.M.P. conditions with initial condition $\hat{x}_i$:

$$\hat{x}_i \longrightarrow (u_i^*(t), x_i^*(t), \lambda_i^*(t)), \quad t \in [0, T]$$

   where x_i^* is originating from $x(0) = \hat{x}_i$
3. Construct a **synthetic dataset**:

$$\begin{cases} V_i := V(\hat{x}_i, 0) = \int_0^T \ell(x_i^*(t)) + \beta \cdot ||u_i^*(t)||^2 dt \\ \nabla V_i := \nabla V(\hat{x}_i, 0) = \lambda_i^*(0) \end{cases}$$

4. Use **supervised learning** (regression) or **unsupervised learning** to obtain $V(x)$ (equivalent to learning a model for $V(x, 0) := V(x)$):
   In the supervised case we can proceed in different ways

- **Neural Networks**:
  We have
  $$V(x) \approx V_\sigma(x) := NN(x, \theta, \ell)$$
  where θ are the **parameters** and ℓ are the **layers**. The **training data** is:
  $$\{\hat{x_i}, V_i, \nabla V_i\}_{i=1}^{N_s}$$
  while the **loss function** is:
  $$\ell(\theta) := \sum_{i=1}^{N_s} ||V_i - V_\theta(\hat{x_i})||^2$$
  In the **gradient augmented** case the loss function becomes:
  $$\ell(\theta) := \sum_{i=1}^{N_s} ||V_i - V_\theta(\hat{x_i})||^2 + \gamma \cdot ||V_i - V_\theta(\hat{x_i})||^2$$
  Learning a model is equivalent to solving $\theta^* = \arg\min_\theta \ell(\theta)$, however this is a **large scale, non-convex** optimization problem!
- **Polynomial approximation**:
  We have
  $$V(x) \approx V_\sigma(x) := \sum_{i=1}^{m} \theta_i \cdot \Phi_i$$
  where $\{\Phi_i\}_{i=1}^m$ are a **polynomial basis** for functions $\mathbb{R}^n \longrightarrow \mathbb{R}$ (e.g. monomial basis, orthogonal polynomials (Legendre/Hermite etc. ...) ...)
  Here we have the advantage that the gradient of V_σ can be trivially computed:
  $$\nabla V_\sigma = \sum_{i=1}^{m} \theta_i \cdot \nabla \Phi_i$$
  Moreover, training this model is equivalent to solving a **linear least squares** problem:
  $$V_\sigma(\hat{x_i}) := \sum_{i=1}^{m} \theta_i \cdot \Phi_i(\hat{x_i}) =< \theta, \Phi(\hat{x_i}) >$$
  $$\implies \ell(\theta) := ||A\theta - (V_i, \nabla V_i, ...)^T||^2, \quad A = \begin{pmatrix} \Phi_1(\hat{x_1}) & \Phi_1(\hat{x_2}) & ... \\ \Phi_2(\hat{x_1}) & \Phi_2(\hat{x_2}) & ... \\ ... & ... & ... \end{pmatrix}$$
  where we note that A is a **Vandermonde Matrix**. We can also use a **sparsity promoting** loss function:
  $$\ell(\theta) = ||A\theta - \overline{V}||^2 + \gamma \cdot ||\theta||_1$$
  where $\overline{V}$ is the **dataset** and the term $\gamma \cdot ||\theta||_1$ is the **sparsity/penalization** term. Notice that this problem is still **convex** but it is **non-smooth**! This means that we need to adopt a Lasso/Proximal Gradient method to solve it. We also need to think how to scale this problem properly in high dimensional cases. If $\dot{x} = Ax + Bu, \quad \ell(x) = x^T Q x$ with $Q \geq 0$, then $V(x) = x^T \Pi x$ (**linear-quadratic control**, where $\Pi \geq 0$ solves a Riccati equation) can be retrieved by polynomial approximation.

- **Physics Informed Neural Networks/Deep Galerkin Method**
  In this case we are in an unsupervised setting and we have:

  $$V(x) \approx V_\theta(x) = NN(x, \theta)$$

  The Hamilton-Jacobi-Bellman PDE is:

  $$||\nabla V|| = 1 \implies ||\nabla V|| - 1 = 0 \implies \text{Res}(V) := ||\nabla V|| - 1$$

  Therefore, solving the PDE is equivalent to solving $\text{Res}(V) = 0$. We introduce the **residual loss**:

  $$\ell(\theta) := ||\text{Res}(V_\theta)||^2_{L^2(\Omega)}$$

  Finally, training this model is equivalent to solving the following optimization problem:

  $$\min_\theta \{|| \, ||\nabla V_\theta|| - 1||^2_{L^2(\Omega)}\}$$