Stochastic Optimal Control under Constraints
with Deep Learning

Kim Anh PHAM

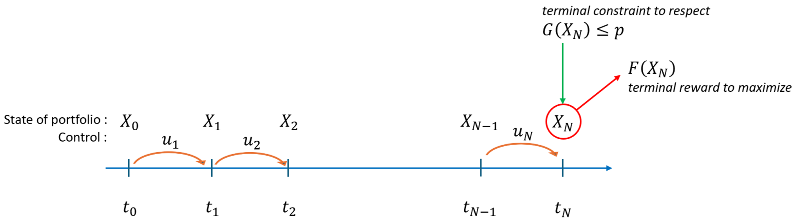
CEREMADE - Paris Dauphine PSL

Young Researcher Days 03-05.06.2025

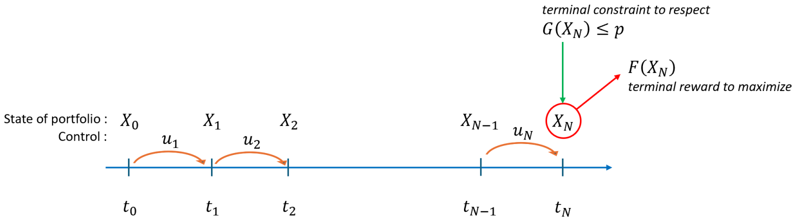
Outline

Portfolio Management basis visualized

State process $X = (X_t)_{t_0 \leq t \leq t_N}$ takes values in $\mathbf{R}^d$.
Control process $u = (u_t)_{t_0 < t \leq t_N}$ takes values in $U \subset \mathbf{R}^{d'}$ compact.



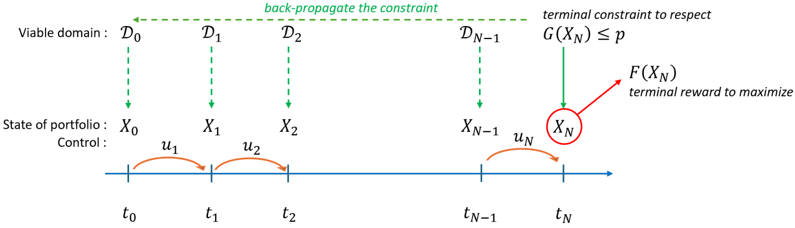
Portfolio Management basis visualized



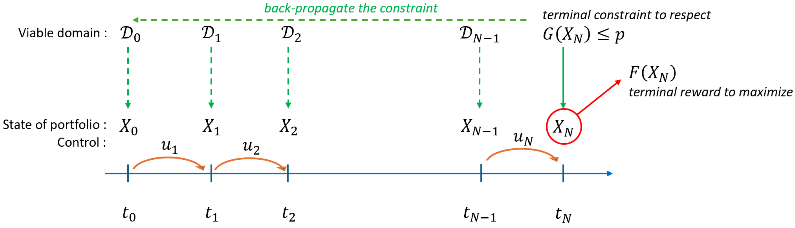
$$\Delta X^u_{t_0,x_0}(t_i) = \mu\left(X^u_{t_0,x_0}(t_i), u_t\right)\Delta t + \sigma\left(X^u_{t_0,x_0}(t_i), u_t\right)\Delta W_{t_i}$$
$$X^u_{t_0,x_0}(t_0) = x_0$$

with $\Delta W_{t_i} := W_{t_i} - W_{t_{i-1}}$ multi-dimensional Brownian increment.

Portfolio Management basis visualized



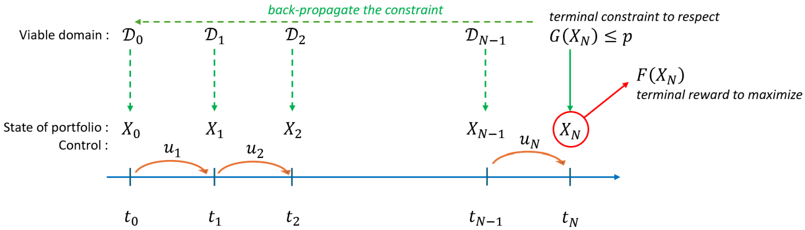
Portfolio Management basis visualized



$$\mathcal{D} := \{(t, x, p) \in [t_0, t_N] \times \mathbf{R}^{d+1} : \exists u \text{ admissible s.t } G(X^u_{t_0,x_0}(t_N) \leq p\}$$

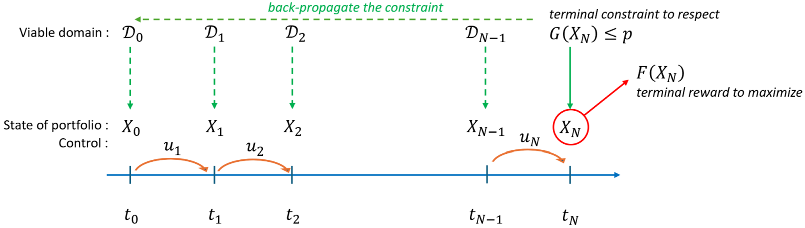
$$\mathcal{U}(t_0, x_0, p) := \{u = (u_{t_i})_{i=1,\dots N} \text{ admissible } : G(X^u_{t_0,x_0}(t_N)) \leq p\}$$

Portfolio Management basis visualized



Value function : $V(t_0, x_0, p) := \max_{u \in \mathcal{U}(t_0, x_0, p)} \mathbb{E}\left[F(X^u_{t_0,x_0}(t_N))\right]$

Portfolio Management basis visualized



Value function : $V(t_0, x_0, p) := \max_{u \in \mathcal{U}(t_0, x_0, p)} \mathbb{E}\left[F(X^u_{t_0,x_0}(t_N))\right]$

Questions :

- What are the characteristics of V ? (PDE)
- Existence ? Unicity ? Smoothness ?
- Numerically, how to solve this problem ?

Continuous framework

For $t \in [0, T], x \in \mathbf{R}^d, u \in \mathcal{U}$ (the set of admissible control),

$$X^u_{t,x}(s) = x + \int_t^s \mu(X^u_{t,x}(r), u_r)dr + \int_t^s \sigma((X^u_{t,x}(r), u_r)dr) \forall t \leq s \leq T \tag{1}$$

where μ and σ are bounded, continuous, and Lipschitz in its first variable uniformly in the second one.
For a given $p \in \mathbf{R}$, the value function

$$V(t, x, p) := \sup_{u \in \mathcal{U}(t,x,p)} \mathbb{E}[F(X^u_{t,x}(T))] \tag{2}$$

where
$\mathcal{U}(t, x, p) := \{u \in \mathcal{U} : \mathbb{E}\left[G(X^u_{t,x}(T)) + \int_t^T g(X^u_{t,x}(r), u_r)dr\right] \leq p\}$
and F, G, and g are continuous with polynomial growth.

Viable Domain and its characterization

We define the viable domain $\mathcal{D}$

$$\begin{aligned} \mathcal{D} &:= \{(t,x,p) \in [0,T] \times \mathbf{R}^{d+1} : \mathcal{U}(t,x,p) \neq \emptyset\} \\ &= \{(t,x,p) \in [0,T] \times \mathbf{R}^{d+1} : \exists u \in \mathcal{U} \text{ s.t} \\ &\quad \mathbb{E}\left[G(X^u_{t,x}(T)) + \int_t^T g(X^u_{t,x}(r), u_r)\right] \leq p\} \end{aligned} \tag{3}$$

Viable Domain and its characterization

We define the viable domain $\mathcal{D}$

$$\begin{aligned}\mathcal{D} &:= \{(t,x,p) \in [0,T] \times \mathbf{R}^{d+1} : \mathcal{U}(t,x,p) \neq \emptyset\} \\ &= \{(t,x,p) \in [0,T] \times \mathbf{R}^{d+1} : \exists u \in \mathcal{U} \text{ s.t} \\ &\quad \mathbb{E}\left[G(X^u_{t,x}(T)) + \int_t^T g(X^u_{t,x}(r), u_r)\right] \leq p\} \end{aligned} \tag{3}$$

Alternatively, we can also look at the boundary value of the viable domain defined as

$$\begin{aligned} w(t,x) &:= \inf\{p : \mathcal{U}(t,x,p) \neq \emptyset\} \\ &= \inf_{u \in \mathcal{U}} \mathbb{E}\left[G(X^u_{t,x}(T)) + \int_t^T g(X^u_{t,x}(r), u_r)\right] \end{aligned} \tag{4}$$

Then, we can write $\mathcal{D} = \{(t,x,p) \in [0,T] \times \mathbf{R}^{d+1} : p \geq w(t,x)\}$.

Viable Domain and its characterization

Furthermore, we can partition the viable domain as:

$$int_P\mathcal{D} := \{(t,x,p) \in [0,T) \times \mathbf{R}^{d+1} : p > w(t,x)\}$$
$$\partial_P\mathcal{D} := \{(t,x,p) \in [0,T) \times \mathbf{R}^{d+1} : p = w(t,x)\}$$
$$\partial_T\mathcal{D} := \{(t,x,p) \in \{T\} \times \mathbf{R}^{d+1} : p \geq w(t,x)\}$$

Viable Domain and its characterization

Furthermore, we can partition the viable domain as:

$$int_P \mathcal{D} := \{(t,x,p) \in [0,T) \times \mathbf{R}^{d+1} : p > w(t,x)\}$$
$$\partial_P \mathcal{D} := \{(t,x,p) \in [0,T) \times \mathbf{R}^{d+1} : p = w(t,x)\}$$
$$\partial_T \mathcal{D} := \{(t,x,p) \in \{T\} \times \mathbf{R}^{d+1} : p \geq w(t,x)\}$$

Why this partition ?

Martingale representation of the constraint

Key idea : represent the constraint value p by a Martingale process.

$$\mathbb{E}\left[G(X^u_{t,x}(T))\right] \leq p \iff \exists a \in \mathcal{A} \text{ s.t } M^a_{t,p} \geq w(\cdot, X^u_{t,x}) \text{ on } [t, T]$$

with

$$M^a_{t,p} := p + \int_t^{\cdot} a_r dW_r$$

where $\mathcal{A}$ is a set of progressively measurable processes such that $M^a_{t,0}$ is a martingale.

Martingale representation of the constraint

$$M^a_{t,p} := p + \int_t^\cdot a_r dW_r$$

Then, the boundary (value) of the viable domain becomes

$$w(t,x) = \inf\{p \in \mathbf{R} : \exists (u,a) \in \mathcal{U} \times \mathcal{A} \text{ s.t } M^a_{t,p} \geq w(\cdot, X^u_{t,x}) \text{ on } [t, T]\}$$

and the value function V can be rewritten as

$$V(t,x,p) = \sup\{\mathbb{E}\left[F(X^u_{t,x}(T))\right] : \\ (u,a) \in \mathcal{U} \times \mathcal{A} \text{ s.t } M^a_{t,p} \geq w(\cdot, X^u_{t,x}) \text{ on } [t, T]\}$$

Further discussion : Bouchard et al. [1], Bouchard et al. [2]

PDE Characterization - Operator and Envelopes

Consider the operator

$$H(t,x,q,q',A) := -\sup_{(u,a)\in U\times\mathbb{R}^d} \bar{\mu}(x,u)^\top q + \frac{1}{2}\mathrm{Tr}\left[(\bar{\sigma}\bar{\sigma}^\top)(x,u,a)A\right] + g(x,u)q'$$

defined for $(t,x,q,q',A) \in [0,T] \times \mathbf{R}^d \times \mathbf{R} \times \mathbf{R} \times \mathbf{S}^{d+1}$ where

$$\bar{\mu}(\cdot,u) := \begin{pmatrix} \mu(\cdot,u) \\ 0 \end{pmatrix} \text{ and } \bar{\sigma}(\cdot,u,a) := \begin{bmatrix} \sigma(\cdot,u) \\ a \end{bmatrix}, (u,a) \in U\times\mathbf{R}^d$$

Let H^* and H_* be upper- and lower-semicontinuous envelopes of H. Similarly, V^* and V_* are upper- and lower-semicontinuous envelopes of V.

PDE Characterization - Hamiltonian-Jacobi-Bellman

Theorem - Bouchard et al, [2]

V_* is a viscosity super-solution of

$$-\partial_t \phi + H^*(t, x, D_{(x,p)}\phi, -D_p\phi, D^2_{(x,p)}\phi) \geq 0 \text{ on } int_p\mathcal{D} \quad (5)$$

$$\phi(T, \cdot) \geq F \text{ on } \{(x,p) \in \mathbf{R}^{d+1} : p > G(x)\} \quad (6)$$

and V^* is a viscosity sub-solution of

$$-\partial_t \phi + H^*(t, x, D_{(x,p)}\phi, -D_p\phi, D^2_{(x,p)}\phi) \leq 0 \text{ on } int_p\mathcal{D} \quad (7)$$

$$\phi(T, \cdot) \leq F \text{ on } \{(x,p) \in \mathbf{R}^{d+1} : p \geq G(x)\} \quad (8)$$

PDE Characterization - Boundary of the domain

Theorem - Bouchard et al, upcoming

Assume that

1. $G \in C_b^2(\mathbf{R}^d)$, g is bounded, and $D^2 G$ and $g(\cdot, u)$ are Hölder continous on $\mathbf{R}^d$, uniformly in $u \in U$. (A)
2. There exists $0 \leq \lambda_\sigma \leq \Lambda_\sigma$ such that $\lambda_\sigma \leq z^\top(\sigma\sigma^\top)(x,u,z) \leq \Lambda_\sigma \forall (x,u,z) \in \mathbf{R}^d \times U \times \partial B_1$ (B)

Then w is a smooth solution to

$$0 = -\inf_{u \in U} \mathcal{L}_X^u w + g \text{ on } [0,T) \times \mathbf{R}^d \tag{9}$$

where $\mathcal{L}_X^u \phi := \partial_t \phi + \mu(\cdot,u)^\top D\phi + \frac{1}{2}\mathrm{Tr}\left[(\sigma\sigma^\top)(\cdot,u) D^2\phi\right]$ is the Dynkin operator for any smooth function ϕ and $u \in U$.

PDE Characterization - Boundary of the domain

Assumption

Let $U(t,x) := \arg\min_{u \in U}\{\mathcal{L}_X^u w + g\}$ be the set of optimal control at a given coordinate $(t,x) \in [0,T) \times \mathbf{R}^d$. We assume that $U(t,x)$ is non-empty and 'continuous' for every $(t,x) \in [0,T) \times \mathbf{R}^d$. (C)

PDE Characterization - Boundary of the domain

Assumption

Let $U(t,x) := \arg\min_{u\in U}\{\mathcal{L}_X^u w + g\}$ be the set of optimal control at a given coordinate $(t,x) \in [0,T) \times \mathbf{R}^d$. We assume that $U(t,x)$ is non-empty and 'continuous' for every $(t,x) \in [0,T) \times \mathbf{R}^d$. (C)

Consider $\mathcal{V} := V(\cdot, w(\cdot))$ and its envelopes

$$\mathcal{V}_*(t,x) := \liminf\{V(t',x',w(t',x')) : [0,T] \times \mathbf{R}^d \ni (t',x') \to (t,x)\}$$

$$\mathcal{V}^*(t,x) := \limsup\{V(t',x',w(t',x')+\varepsilon) : [0,T] \times \mathbf{R}^d \times (0,\infty) \ni (t',x',\varepsilon) \to (t,x,0)\}$$

PDE Characterization - Boundary of the domain

Assumption

Let $U(t,x) := \arg\min_{u \in U}\{\mathcal{L}_X^u w + g\}$ be the set of optimal control at a given coordinate $(t,x) \in [0,T) \times \mathbf{R}^d$. We assume that $U(t,x)$ is non-empty and 'continuous' for every $(t,x) \in [0,T) \times \mathbf{R}^d$. (C)

Theorem - Bouchard et al, upcoming

Given assumptions (A) and (B), $\mathcal{V}^*$ is a viscosity sub-solution of

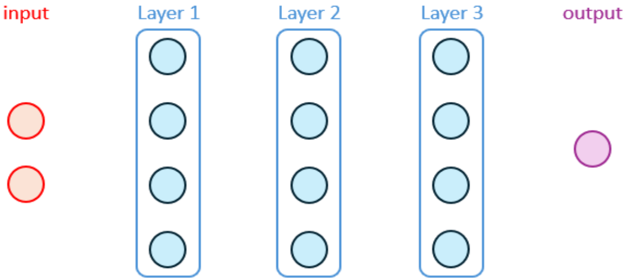
$$-\max_{u \in U(t,x)} \mathcal{L}_X^u \phi(t,x) = 0, \forall (t,x) \in [0,T) \times \mathbf{R}^d \quad (10)$$

$$\phi(T,x) = F(x), \forall x \in \mathbf{R}^d \quad (11)$$

Moreover, if assumption (C) also holds, then $\mathcal{V}_*$ is a viscosity super-solution of (10)-(11)

Introduction to Neural Network

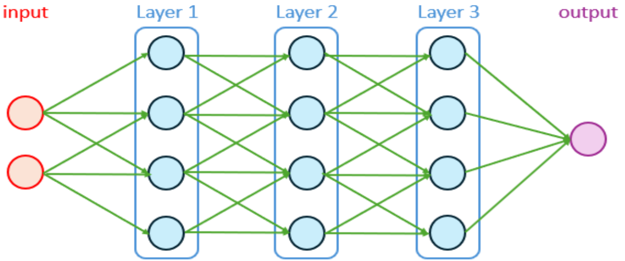
Basic layout of a neural network



Hyper-parameters : dimension of input, number of layers, number of neuron per layer (assuming fully connected), dimension of output

Introduction to Neural Network

Mathematical basis of a neural network



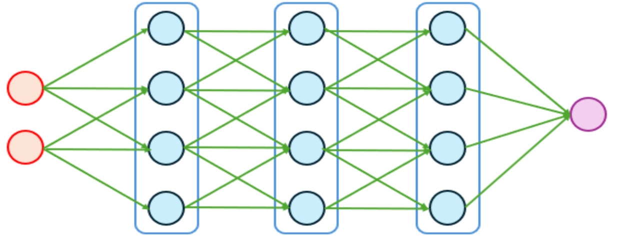
Given the i-th layer of the network with an activation function $\mathcal{F}$ and the input $\mathcal{I}_i$, the output of this layer is

$$\mathcal{O}_i^{\theta} = \mathcal{F}(W_i \mathcal{I}_i + b_i)$$

where W_i is the weight and b_i is the bias.

Introduction to Neural Network

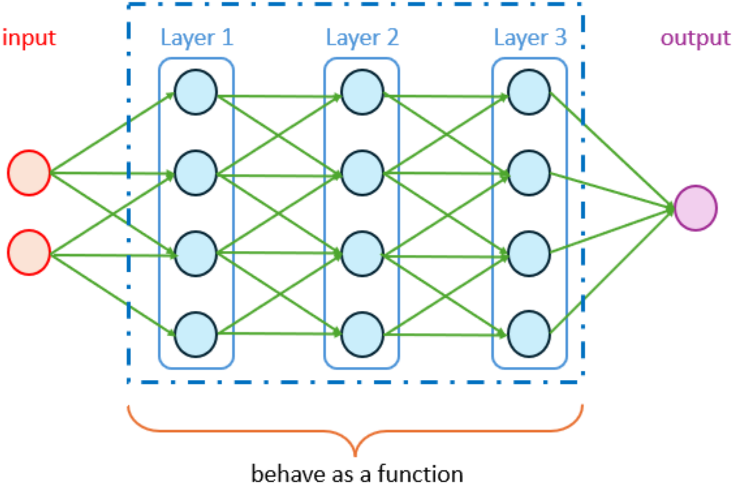
Mathematical basis of a neural network



Final output :

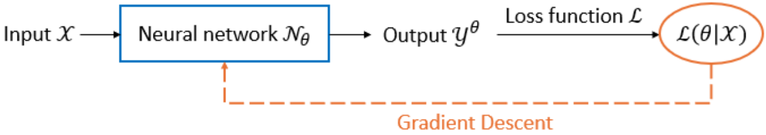
$$
\begin{aligned}
\mathcal{O}^\theta &= \mathcal{F}(W_3 \cdot \mathcal{O}_2^\theta + b_3) \\
&= \mathcal{F}(W_3 \cdot \mathcal{F}(W_2 \cdot \mathcal{O}_1^\theta + b_2) + b_3) \\
&= \mathcal{F}(W_3 \cdot \mathcal{F}(W_2 \cdot \mathcal{F}(W_1 \cdot \mathcal{I} + b_1) + b_2) + b_3)
\end{aligned}
$$

Introduction to Neural Network



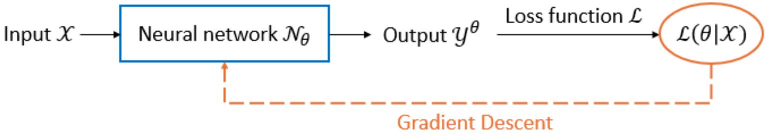
Training a neural network

Pass forward process :



Training a neural network

Pass forward process :



Key idea for training :

- pass forward multiple times
- update parameters after each pass (to lower loss)
- stop the loop when the loss is as low as possible/desirable

Why Neural Networks ?

Classical Method

ex : Finite Difference Method

Prone to discretization error

+

Curse of Dimensionality

Deep Learning

ex : Physic-Informed Neural Networks

Can handle high dimension

+

Fast computation / High precision

Why Neural Networks ?

Promising, yet more research is needed !

Classical Method	**Deep Learning**
ex : Finite Difference Method	ex : Physic-Informed Neural Networks
Prone to discretization error	Can handle high dimension
+	+
Curse of Dimensionality	Fast computation / High precision

Proposed Algorithm

Numerical resolution steps

1. Estimate the optimal control process $u^\theta = (u^\theta_{t_i})_{i=1,\dots N}$
2. Estimate the boundary value of the viable domain w_θ
3. Estimate the value function at the boundary $\mathcal{V}_\theta$
4. Estimate the optimal control process with the martingal increment $(\tilde{u}, \tilde{a})_\theta$
5. Estimate the value function on the entire viable domain V_θ

Step 1 : Estimate the optimal process u^θ

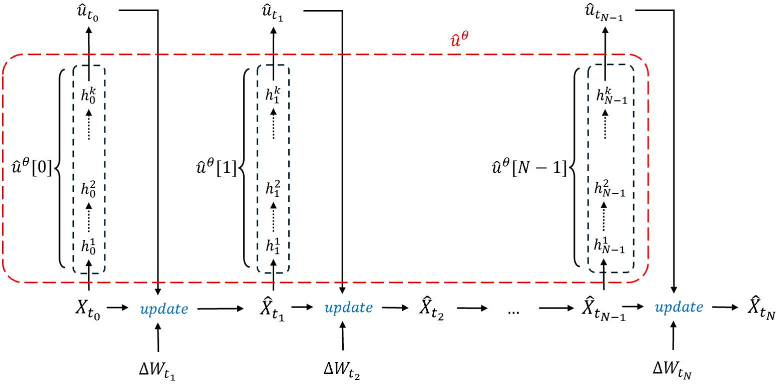
Recall the definition of the optimal control process u^* (discretized)

$$u^* = (u^*_{t_i})_{i=1,\dots N} = \arg\min \mathbb{E}\left[G(X_{t_N}) + \frac{1}{N}\sum_{i=1}^{N} g(X_{t_i})\right]$$

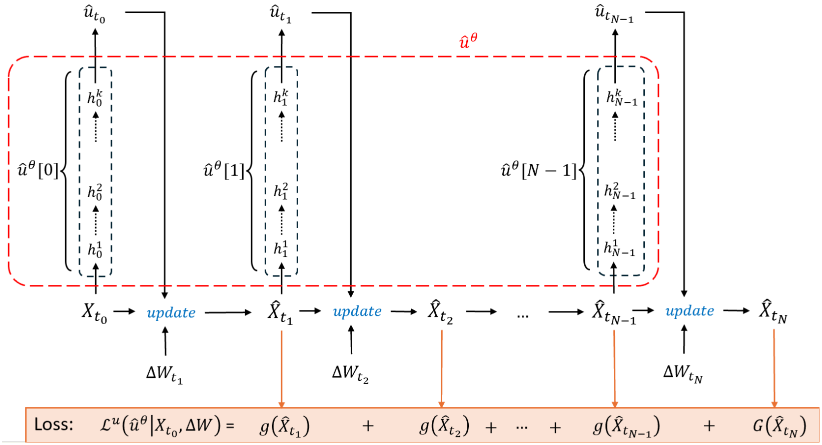
Given a sample of initial states $\mathcal{X}_0 = (X^j_{t_0})^{j=1,\dots J}$ and a sample of brownian increments $\Delta\mathcal{W} = ((\Delta W_{t_i})_{i=1,\dots N})^{j=1,\dots J}$, practically we seek to minizie the empirical mean

$$\hat{u}^\theta = (\hat{u}^\theta_{t_i})_{i=1,\dots N} = \arg\min \sum\nolimits_{j=1}^{J}\left[G(\hat{X}^j_{t_N}) + \frac{1}{N}\sum_{i=1}^{N} g(\hat{X}^j_{t_i})\right]$$

Step 1 : Estimate the optimal process u^θ



Step 1 : Estimate the optimal process u^θ



Step 1 : Estimate the optimal process u^θ

Pseu-do code :

1. Randomly generate a sample of initial states $\mathcal{X}_0 = (X^j_{t_0})^{j=1,...J}$ and a sample of brownian increments $\Delta\mathcal{W} = ((\Delta W_{t_i})_{i=1,...N})^{j=1,...J}$
2. For $e = 1$ to N^u_e ($=$ number of epoch):
   - Pass forward $(\mathcal{X}_0, \Delta\mathcal{W})$ through the network $\hat{u}^\theta$ to get the full trajectories $\hat{\mathcal{X}} = \left((\hat{X}^j_{t_i})_{i=1,..N}\right)^{j=1,...J}$
   - Compute the cumulative loss $\mathcal{L}^u(\hat{u}^\theta)$
   - Take a gradient descent step

Step 2: Estimate the boundary value w_θ

In theory, w should satisfy

$$0 = -\inf_{u\in U} \mathcal{L}_X^u w + g \text{ on } [0, T) \times \mathbf{R}^d$$
$$w(T, \cdot) = G \text{ on } \mathbf{R}^d$$

Assuming that the optimal control network $\hat{u}^\theta$ is well trained, then within our discretized framework, we would train $\hat{w}_\theta$ to satisfy

$$0 = \mathcal{L}_X^{\hat{u}_{t_i}} \hat{w}_\theta + g \forall t_i \in \{t_0, \dots t_{N-1}\}$$
$$\hat{w}_\theta(t_N, \hat{X}_{t_N}) = G(\hat{X}_{t_N})$$

where $(\hat{X}_{t_i})_{i=1,\dots N}$ is any trajectory generated by passing a randomly selected initial state x_0, a random Brownian movement W through the trained network $\hat{u}^\theta$.

Step 2: Estimate the boundary value w_θ

Physics-Informed Neural Network : a method used to train neural network which emphasizes on the (PDE) characteristics of the function that the network aims to approximate.

Step 2: Estimate the boundary value w_θ

Physics-Informed Neural Network : a method used to train neural network which emphasizes on the (PDE) characteristics of the function that the network aims to approximate.

Data for training :

$$(\mathcal{X}_0, \Delta\mathcal{W}) = \left(x_0^j, \left(\Delta W_{t_i}\right)_{i=1,\dots N}\right)^{j=1,..J} \xrightarrow{\hat{u}^\theta} \left(\left(\hat{X}_{t_i}^j\right)_{i=1,\dots N}\right)^{j=1,..J} = \left(\left(X_{t_i}^{\hat{u}^\theta, t_0, x_0^j}\right)_{i=1,..N}\right)^{j=1,..J}$$

Step 2: Estimate the boundary value w_θ

Physics-Informed Neural Network : a method used to train neural network which emphasizes on the (PDE) characteristics of the function that the network aims to approximate.

Data for training :

$$(\mathcal{X}_0, \Delta\mathcal{W}) = \left(x_0^j, \left(\Delta W_{t_i}\right)_{i=1,\dots N}\right)^{j=1,..J} \xrightarrow{\hat{u}^\theta} \left(\left(\hat{X}_{t_i}^j\right)_{i=1,\dots N}\right)^{j=1,..J} = \left(\left(X_{t_i}^{\hat{u}^\theta, t_0, x_0^j}\right)_{i=1,..N}\right)^{j=1,..J}$$

Loss function :

$$\mathcal{L}^w(w_\theta|\mathcal{X}_0, \Delta\mathcal{W}) := \frac{1}{J}\sum_{j=1}^{J}\frac{1}{N}\sum_{i=0}^{N-1}\left|\mathcal{L}_X^u w_\theta(t_i, \hat{X}_{t_i}^j) + g(\hat{X}_{t_i}^j, \hat{u}_{t_i}^{\theta,j})\right|^2 + \frac{1}{J}\sum_{j=1}^{J}\left|w_\theta(t_N, \hat{X}_{t_N}^j) - G(\hat{X}_{t_N}^j)\right|^2$$

Thank you for your attention !

Bibliography

Bruno Bouchard, Romuald Elie, and Nizar Touzi.
Stochastic target problems with controlled loss.
SIAM Journal on Control and Optimization, 48(5):3123–3150, 2010.

Bruno Bouchard and Marcel Nutz.
Weak dynamic programming for generalized state constraints.
SIAM Journal on Control and Optimization, 50(6):3344–3373, January 2012.