Deep Leaning and Partial Differntial Equation

Liyao Lyu

Neural Network

Fully connected neural networks

Neural Network

Fully connected neural networks

- Input: x or t, x

Neural Network

Fully connected neural networks

- Input: x or t, x

- Output: $\tilde{u}(x; \theta)$ or $\tilde{u}(t, x; \theta)$

Neural Network

Fully connected neural networks

- Input: x or t, x

- Output: $\tilde{u}(x;\theta)$ or $\tilde{u}(t,x;\theta)$

- $s^1 = \sigma(W^1 \cdot x + b^1)$
  $s^k = \sigma(W^k \cdot s^{k-1} + b^k)$ for $k = 2, \cdots, L$
  $\tilde{u}(x;\theta) = W^{L+1} \cdot s^L + b^L$

Neural Network

Fully connected neural networks

- Input: x or t, x

- Output: $\tilde{u}(x;\theta)$ or $\tilde{u}(t,x;\theta)$

- $s^1 = \sigma(W^1 \cdot x + b^1)$
  $s^k = \sigma(W^k \cdot s^{k-1} + b^k)$ for $k = 2,\cdots,L$
  $\tilde{u}(x;\theta) = W^{L+1} \cdot s^L + b^L$

- $\theta = \{W^1, b^1, \cdots, W^L, b^L\}$ are paremeters to be optimized and σ is the activation function

PINN or DGM

PINN or DGM

- Our Object is to solve the Partial Differential Equation (PDE)

$$\begin{cases} \partial_t u + \mathscr{L}[u] = 0 & (t, x) \in [0,T] \times \Omega \\ u(0,t) = g(x) & x \in \Omega \\ u(t, x) = f(t, x) \qquad (t, x) & (t, x) \in [0,T] \times \partial\Omega \end{cases}$$

PINN or DGM

- Our Object is to solve the Partial Differential Equation (PDE)

$$\begin{cases} \partial_t u + \mathscr{L}[u] = 0 & (t,x) \in [0,T] \times \Omega \\ u(0,t) = g(x) & x \in \Omega \\ u(t,x) = f(t,x) \qquad (t,x) & (t,x) \in [0,T] \times \partial\Omega \end{cases}$$

- Define the loss function

$$J[u] = \lambda_1 \|\partial_t u - \mathscr{L}u\|^2_{2,[0,T]\times\Omega} + \lambda_2 \|u(0,x) - g(x)\|^2_{2,\Omega} + \lambda_3 \|u(t,x) - f(t,x)\|^2_{2,[0,T]\times\partial\Omega}$$

PINN or DGM

- Our Object is to solve the Partial Differential Equation (PDE)

$$\begin{cases} \partial_t u + \mathscr{L}[u] = 0 & (t,x) \in [0,T] \times \Omega \\ u(0,t) = g(x) & x \in \Omega \\ u(t,x) = f(t,x) \qquad (t,x) & (t,x) \in [0,T] \times \partial\Omega \end{cases}$$

- Define the loss function

$$J[u] = \lambda_1 \|\partial_t u - \mathscr{L}u\|^2_{2,[0,T]\times\Omega} + \lambda_2 \|u(0,x) - g(x)\|^2_{2,\Omega} + \lambda_3 \|u(t,x) - f(t,x)\|^2_{2,[0,T]\times\partial\Omega}$$

- Solve the PDE by minimize the loss function $\arg\min_{\theta} J[\tilde{u}[t,x;\theta]]$

Example

Example

- Consider a elliptic partial differential
$$-\sum_{i=1}^{d} \partial_{x_i}\left(\partial_{x_j} a_{ij} u\right) + \sum_{i=1}^{d} b_i \partial_{x_i} u + cu - f = 0$$
with boundary condition $u(x) = g(x)$ on $\partial\Omega$

Example

- Consider a elliptic partial differential
$$-\sum_{i=1}^{d} \partial_{x_i}\left(\partial_{x_j} a_{ij} u\right) + \sum_{i=1}^{d} b_i \partial_{x_i} u + cu - f = 0$$
with boundary condition $u(x) = g(x)$ on $\partial\Omega$

- Then
$$\arg\min_{\theta} J[\tilde{u}(x;\theta)] = \int_{\Omega}\left(-\sum_{i=1}^{d} \partial_{x_i}\left(\partial_{x_j} a_{ij}\tilde{u}\right) + \sum_{i=1}^{d} b_i \partial_{x_i}\tilde{u} + c\tilde{u} - f\right)^2 dx + \int_{\partial\Omega} (\tilde{u}(x) - g(x))^2 dx$$
with gradient is computed by autogradient and integral is compute by Monte Carlo sampling in Ω and $\partial\Omega$

Example

Admission set

- $\arg\min_{\theta} J[\tilde{u}(x;\theta)] = \int_{\Omega} \left(-\sum_{i=1}^{d} \partial_{x_i} \left(\partial_{x_j} a_{ij} \tilde{u} \right) + \sum_{i=1}^{d} b_i \partial_{x_i} \tilde{u} + c\tilde{u} - f \right)^2 dx + \int_{\partial\Omega} (\tilde{u}(x) - g(x))^2 dx$
  with gradient is computed by autogradient and integral is compute by Monte Carlo sampling in Ω and $\partial\Omega$

Example

Admission set

- $\arg\min_{\theta} J[\tilde{u}(x;\theta)] = \int_{\Omega} \left(-\sum_{i=1}^{d} \partial_{x_i} \left(\partial_{x_j} a_{ij} \tilde{u} \right) + \sum_{i=1}^{d} b_i \partial_{x_i} \tilde{u} + c\tilde{u} - f \right)^2 dx + \int_{\partial\Omega} (\tilde{u}(x) - g(x))^2 dx$ with gradient is computed by autogradient and integral is compute by Monte Carlo sampling in Ω and $\partial\Omega$

- Question: Can we remove the boundary condition constraint by design a admissible set?

Example

Admission set

- $\arg\min_{\theta} J[\tilde{u}(x;\theta)] = \int_{\Omega} \left(-\sum_{i=1}^{d} \partial_{x_i} \left(\partial_{x_j} a_{ij} \tilde{u} \right) + \sum_{i=1}^{d} b_i \partial_{x_i} \tilde{u} + c\tilde{u} - f \right)^2 dx + \int_{\partial\Omega} (\tilde{u}(x) - g(x))^2 dx$ with gradient is computed by autogradient and integral is compute by Monte Carlo sampling in Ω and $\partial\Omega$

- Question: Can we remove the boundary condition constraint by design a admissible set?

- Answer: We can try to construct NN such that automatically satisfies the boundry condition, for example $\tilde{u}(x;\theta) = D(x)\mathcal{N}(x;\theta) + g(x)$

Example

Weak formula and adversarial training

- $-\sum_{i=1}^{d} \partial_{x_i} \left(\partial_{x_j} a_{ij} u \right) + \sum_{i=1}^{d} b_i \partial_{x_i} u + cu - f = 0$

Example

Weak formula and adversarial training

- $-\sum_{i=1}^{d} \partial_{x_i}\left(\partial_{x_j} a_{ij} u\right) + \sum_{i=1}^{d} b_i \partial_{x_i} u + cu - f = 0$
- Take the weak formula $\forall \varphi(x)$, we multiply φ on the both side, we have $-\sum_{i=1}^{d} \partial_{x_i}\left(\partial_{x_j} a_{ij} u\right)\varphi + \sum_{i=1}^{d} b_i \partial_{x_i} u\varphi + cu\varphi - f\varphi = 0$ for any φ.

Example

Weak formula and adversarial training

- $-\sum_{i=1}^{d} \partial_{x_i} \left(\partial_{x_j} a_{ij} u \right) + \sum_{i=1}^{d} b_i \partial_{x_i} u + cu - f = 0$

- Take the weak formula $\forall \varphi(x)$, we multiply φ on the both side, we have
  $-\sum_{i=1}^{d} \partial_{x_i} \left(\partial_{x_j} a_{ij} u \right) \varphi + \sum_{i=1}^{d} b_i \partial_{x_i} u \varphi + cu\varphi - f\varphi = 0$ for any φ.

- The loss function can be writen as
  $$\arg\max_{\theta_\varphi} \arg\min_{\theta_u} J[\tilde{u}(x;\theta)] = \int_\Omega \left(\sum_{i=1}^{d} a_{ij} \partial_{x_i} \tilde{u} \partial_{x_j} \tilde{\varphi} + \sum_{i=1}^{d} b_i \partial_{x_i} \tilde{u} \tilde{\varphi} + c\tilde{u}\tilde{\varphi} - f\tilde{\varphi} \right)^2 dx.$$

Example

Weak formula and adversarial training

Example

Weak formula and adversarial training

- $\arg\max_{\theta_\varphi} \arg\min_{\theta_u} J[\tilde{u}(x;\theta)] = \int_\Omega \left(\sum_{i=1}^{d} a_{ij}\partial_{x_i}\tilde{u}\partial_{x_j}\tilde{\varphi} + \sum_{i=1}^{d} b_i\partial_{x_i}\tilde{u}\tilde{\varphi} + c\tilde{u}\tilde{\varphi} - f\tilde{\varphi} \right)^2 dx.$

Example

Weak formula and adversarial training

- $$\arg\max_{\theta_\varphi} \arg\min_{\theta_u} J[\tilde{u}(x;\theta)] = \int_\Omega \left(\sum_{i=1}^{d} a_{ij}\partial_{x_i}\tilde{u}\partial_{x_j}\tilde{\varphi} + \sum_{i=1}^{d} b_i \partial_{x_i}\tilde{u}\tilde{\varphi} + c\tilde{u}\tilde{\varphi} - f\tilde{\varphi} \right)^2 dx.$$

- In pratice, we may add some entropy regularization

$$\arg\max_{\theta_\varphi} \arg\min_{\theta_u} J[\tilde{u}(x;\theta)] = \log \left(\int_\Omega \left(\sum_{i=1}^{d} a_{ij}\partial_{x_i}\tilde{u}\partial_{x_j}\tilde{\varphi} + \sum_{i=1}^{d} b_i \partial_{x_i}\tilde{u}\tilde{\varphi} + c\tilde{u}\tilde{\varphi} - f\tilde{\varphi} \right)^2 dx \right) + \log \left(\int_\Omega \tilde{\varphi}^2 dx \right)$$

.

Deep Ritz method

Deep Ritz method

- Some PDE can has a variational form like $\min_u I[u]$, where $I[u] = \int_\omega \left(\frac{1}{2} |\nabla u(x)|^2 - f(x)u(x) \right)$

Deep Ritz method

- Some PDE can has a variational form like $\min_u I[u]$, where $I[u] = \int_\omega \left(\frac{1}{2} |\nabla u(x)|^2 - f(x)u(x) \right)$
- For example Poisson Equation:

Deep Ritz method

- Some PDE can has a variational form like $\min_u I[u]$, where $I[u] = \int_\omega \left(\frac{1}{2} |\nabla u(x)|^2 - f(x)u(x) \right)$

- For example Poisson Equation:

$$\begin{aligned} -\Delta u &= 1 \qquad & x \in \Omega \\ u(x) &= 0 \qquad & x \in \partial\Omega \end{aligned}$$

Deep Ritz method

- Some PDE can has a variational form like $\min_u I[u]$, where $I[u] = \int_\omega \left(\frac{1}{2} |\nabla u(x)|^2 - f(x)u(x) \right)$
- For example Poisson Equation:

$$
\begin{aligned}
-\Delta u &= 1 & x &\in \Omega \\
u(x) &= 0 & x &\in \partial\Omega
\end{aligned}
$$

$$I[u] = \int_\Omega \left(\frac{1}{2} |\nabla u|^2 - u \right) dx$$

Deep Ritz method

- Some PDE can has a variational form like $\min_{u} I[u]$, where $I[u] = \int_{\omega} \left(\frac{1}{2} |\nabla u(x)|^2 - f(x)u(x) \right)$
- Time independent Schrödinger's Equation under externel potential $V(x)$:

Deep Ritz method

- Some PDE can has a variational form like $\min_u I[u]$, where $I[u] = \int_\omega \left(\frac{1}{2}|\nabla u(x)|^2 - f(x)u(x)\right)$

- Time independent Schrödinger's Equation under externel potential $V(x)$:

$$
\begin{aligned}
&-\Delta u + V \cdot u = \lambda u && x \in \Omega \\
&u(x) = 0 && x \in \partial\Omega
\end{aligned}
$$

Deep Ritz method

- Some PDE can has a variational form like $\min_u I[u]$, where $I[u] = \int_\omega \left(\frac{1}{2}|\nabla u(x)|^2 - f(x)u(x)\right)$
- Time independent Schrödinger's Equation under externel potential $V(x)$:

$$\begin{aligned} -\Delta u + V \cdot u &= \lambda u && x \in \Omega \\ u(x) &= 0 && x \in \partial\Omega \end{aligned}$$

$$\min \frac{\int_\Omega \left(\frac{1}{2}|\nabla u|^2 + Vu^2\right) dx}{\int_\Omega u^2 dx}$$

Deep BSDE

Deep BSDE

- Some PDE can be writen as a control problem, for example a classical linear quadratic Gaussian control problem,

Deep BSDE

- Some PDE can be writen as a control problem, for example a classical linear quadratic Gaussian control problem,

$$dx_t = b(t, x_t, \alpha_t)dt + \sigma(t, x_t, \alpha_t)dW_t,$$

Deep BSDE

- Some PDE can be writen as a control problem, for example a classical linear quadratic Gaussian control problem,

$$dx_t = b(t, x_t, \alpha_t)dt + \sigma(t, x_t, \alpha_t)dW_t,$$

with $t \in [0,T]$ and $x_0 = x$ and with the cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T f(t, x_s, \alpha_s)ds + g(X_T)\right]$.

Deep BSDE

- Some PDE can be writen as a control problem, for example a classical linear quadratic Gaussian control problem,

$$dx_t = b(t, x_t, \alpha_t)dt + \sigma(t, x_t, \alpha_t)dW_t,$$

with $t \in [0,T]$ and $x_0 = x$ and with the cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T f(t, x_s, \alpha_s)ds + g(X_T)\right].$

Let value function defined as $u(t, x) = \inf_{\alpha} \mathbb{E}\left[\int_t^T f(s, x_s, \alpha_s)ds + g(X_T)\Big| x_t = x\right].$

Deep BSDE

- Some PDE can be writen as a control problem, for example a classical linear quadratic Gaussian control problem,

$$dx_t = b(t, x_t, \alpha_t)dt + \sigma(t, x_t, \alpha_t)dW_t,$$

with $t \in [0,T]$ and $x_0 = x$ and with the cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T f(t, x_s, \alpha_s)ds + g(X_T)\right]$.

Let value function defined as $u(t, x) = \inf_{\alpha} \mathbb{E}\left[\int_t^T f(s, x_s, \alpha_s)ds + g(X_T) \,\middle|\, x_t = x\right]$.

The value function will satiesfied the HJB equation

Deep BSDE

- Some PDE can be writen as a control problem, for example a classical linear quadratic Gaussian control problem,

$$dx_t = b(t, x_t, \alpha_t)dt + \sigma(t, x_t, \alpha_t)dW_t,$$

with $t \in [0,T]$ and $x_0 = x$ and with the cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T f(t, x_s, \alpha_s)ds + g(X_T)\right]$.

Let value function defined as $u(t, x) = \inf_{\alpha} \mathbb{E}\left[\int_t^T f(s, x_s, \alpha_s)ds + g(X_T)\,\Big|\,x_t = x\right]$.

The value function will satiesfied the HJB equation

$$\partial_t u(t, x) + \min_{\alpha}\left(H(t, x, \nabla_x u, \mathrm{Hess}_x u, \alpha)\right) = 0$$

Deep BSDE

- Some PDE can be writen as a control problem, for example a classical linear quadratic Gaussian control problem,

$$dx_t = b(t, x_t, \alpha_t)dt + \sigma(t, x_t, \alpha_t)dW_t,$$

with $t \in [0,T]$ and $x_0 = x$ and with the cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T f(t, x_s, \alpha_s)ds + g(X_T)\right]$.

Let value function defined as $u(t,x) = \inf_\alpha \mathbb{E}\left[\int_t^T f(s, x_s, \alpha_s)ds + g(X_T) \,\Big|\, x_t = x\right]$.

The value function will satiesfied the HJB equation

$$\partial_t u(t,x) + \min_\alpha \left(H(t, x, \nabla_x u, \text{Hess}_x u, \alpha)\right) = 0$$

with $u(T,x) = g(x)$ and $H(t,x,p,q,\alpha) = b(t,x,\alpha) \cdot p + \frac{1}{2}\text{Tr}(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^T q) + f(t,x,\alpha)$

Deep BSDE

Informal Derivation

$$dx_t = b(t, x_t, \alpha_t)dt + \sigma(t, x_t, \alpha_t)dW_t$$

$$\mathbb{E}[dW_t] = 0 \quad E[dW_t^i dW_t^j] = \sqrt{dt}\delta_{ij}$$

Deep BSDE

Informal Derivation

$$dx_t = b(t, x_t, \alpha_t)dt + \sigma(t, x_t, \alpha_t)dW_t$$

$$\mathbb{E}[dW_t] = 0 \quad E[dW_t^i dW_t^j] = \sqrt{dt}\delta_{ij}$$

$$
\begin{aligned}
u(t,x) &= \inf_\alpha \mathbb{E}\left[\int_t^T f(s, x_s, \alpha_s)ds + g(X_T) \,\middle|\, x_t = x\right] \\
&= \inf_\alpha \mathbb{E}\left[\int_t^{t+\delta t} f(s, x_s, \alpha_s)ds + \int_{t+\delta t}^T f(s, x_s, \alpha_s)ds + g(X_T) \,\middle|\, x_t = x\right] \\
&= \inf_\alpha \mathbb{E}\left[\int_t^{t+\delta t} f(s, x_s, \alpha_s)ds + u(t+\delta t, x(t+\delta t)) \,\middle|\, x_t = x\right] \\
&= \inf_\alpha \mathbb{E}\left[\int_t^{t+\delta t} f(s, x_s, \alpha_s)ds + u(t,x) + \partial_t u \delta t + \nabla_x u \delta x + \frac{1}{2}\mathrm{Hess}_x u (\delta x)^2 \,\middle|\, x_t = x\right] \\
&= u(t,x) + \partial_t u \delta t + \inf_\alpha \mathbb{E}\left[f(s, x_s, \alpha_s)\delta t + \nabla_x u b(t,x,\alpha)\delta t + \nabla_x u \sigma(t,x,\alpha)dW_t + \frac{1}{2}\mathrm{Hess}_x u(\delta x)^2 \,\middle|\, x_t = x\right] \\
&= u(t,x) + \partial_t u \delta t + \inf_\alpha \mathbb{E}\left[f(s, x_s, \alpha_s)\delta t + \nabla_x u b(t,x,\alpha)\delta t + \frac{1}{2}\mathrm{Hess}_x u(\delta x)^2 \,\middle|\, x_t = x\right] \\
&= u(t,x) + \partial_t u \delta t + \inf_\alpha \mathbb{E}\left[f(s, x_s, \alpha_s)\delta t + \nabla_x \nabla_x u b(t,x,\alpha)\delta t + \frac{1}{2}\mathrm{Hess}_x u(b^T b \delta t^2 + b^T \sigma \delta t dW_t + (\sigma dW_t)^T(\sigma dW_t)) \,\middle|\, x_t = x\right] \\
&= u(t,x) + \partial_t u \delta t + \inf_\alpha \mathbb{E}\left[f(s, x_s, \alpha_s)\delta t + \nabla_x u b(t,x,\alpha)\delta t + \mathrm{Tr}\left(\mathrm{Hess}_x u \sigma^T \sigma\right)\delta t) \,\middle|\, x_t = x\right]
\end{aligned}
$$

Deep BSDE

Example

Deep BSDE

Example

Consider a Linear-quadratic Gaussian control Problem $dx_t = 2\sqrt{\lambda}\alpha_t dt + \sqrt{2}dW_t$, with cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T \|\alpha_s\|^2 ds + g(x_T)\right]$.

Deep BSDE

Example

Consider a Linear-quadratic Gaussian control Problem $dx_t = 2\sqrt{\lambda}\alpha_t dt + \sqrt{2}dW_t$, with cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T \|\alpha_s\|^2 ds + g(x_T)\right]$.

Here $b(t,x,\alpha) = 2\sqrt{\lambda}\alpha, \sigma(t,x,\alpha) = \sqrt{2}, f(t,x,\alpha) = \|\alpha\|^2$, then

$$
\begin{aligned}
H(t,x,\nabla_x u, \mathrm{Hess}_x u, \alpha) &= b(t,x,\alpha)\cdot\nabla_x u + \frac{1}{2}\mathrm{Tr}(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^T \mathrm{Hess}_x u) + f(t,x,\alpha) \\
&= 2\sqrt{\lambda}\alpha\cdot\nabla_x u + \Delta u + \|\alpha\|^2
\end{aligned}
$$

Deep BSDE

Example

Consider a Linear-quadratic Gaussian control Problem $dx_t = 2\sqrt{\lambda}\alpha_t dt + \sqrt{2}dW_t$, with cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T \|\alpha_s\|^2 ds + g(x_T)\right]$.

Here $b(t,x,\alpha) = 2\sqrt{\lambda}\alpha, \sigma(t,x,\alpha) = \sqrt{2}, f(t,x,\alpha) = \|\alpha\|^2$, then

$$\begin{aligned} H(t,x,\nabla_x u, \text{Hess}_x u, \alpha) &= b(t,x,\alpha)\cdot\nabla_x u + \frac{1}{2}\text{Tr}(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^T \text{Hess}_x u) + f(t,x,\alpha) \\ &= 2\sqrt{\lambda}\alpha\cdot\nabla_x u + \Delta u + \|\alpha\|^2 \end{aligned}$$

The HJB equation is given by

Deep BSDE

Example

Consider a Linear-quadratic Gaussian control Problem $dx_t = 2\sqrt{\lambda}\alpha_t dt + \sqrt{2}dW_t$, with cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T \|\alpha_s\|^2 ds + g(x_T)\right]$.

Here $b(t,x,\alpha) = 2\sqrt{\lambda}\alpha, \sigma(t,x,\alpha) = \sqrt{2}, f(t,x,\alpha) = \|\alpha\|^2$, then

$$
\begin{aligned}
H(t,x,\nabla_x u, \mathrm{Hess}_x u, \alpha) &= b(t,x,\alpha)\cdot\nabla_x u + \frac{1}{2}\mathrm{Tr}(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^T\mathrm{Hess}_x u) + f(t,x,\alpha) \\
&= 2\sqrt{\lambda}\alpha\cdot\nabla_x u + \Delta u + \|\alpha\|^2
\end{aligned}
$$

The HJB equation is given by

$$
\partial_t u(t,x) + \min_\alpha \left(H(t,x,\nabla_x u, \mathrm{Hess}_x u, \alpha)\right) = 0,
$$

Deep BSDE

Example

Consider a Linear-quadratic Gaussian control Problem $dx_t = 2\sqrt{\lambda}\alpha_t dt + \sqrt{2}dW_t$, with cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T \|\alpha_s\|^2 ds + g(x_T)\right]$.

Here $b(t,x,\alpha) = 2\sqrt{\lambda}\alpha, \sigma(t,x,\alpha) = \sqrt{2}, f(t,x,\alpha) = \|\alpha\|^2$, then

$$
\begin{aligned}
H(t,x,\nabla_x u, \text{Hess}_x u, \alpha) &= b(t,x,\alpha)\cdot\nabla_x u + \frac{1}{2}\text{Tr}(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^T \text{Hess}_x u) + f(t,x,\alpha) \\
&= 2\sqrt{\lambda}\alpha\cdot\nabla_x u + \Delta u + \|\alpha\|^2
\end{aligned}
$$

The HJB equation is given by

$$
\partial_t u(t,x) + \min_\alpha \left(H(t,x,\nabla_x u, \text{Hess}_x u, \alpha)\right) = 0,
$$

where H takes minima at $\alpha = -\sqrt{\lambda}\,\nabla_x u$ and

Deep BSDE

Example

Consider a Linear-quadratic Gaussian control Problem $dx_t = 2\sqrt{\lambda}\alpha_t dt + \sqrt{2}dW_t$, with cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T \|\alpha_s\|^2 ds + g(x_T)\right]$.

Here $b(t,x,\alpha) = 2\sqrt{\lambda}\alpha, \sigma(t,x,\alpha) = \sqrt{2}, f(t,x,\alpha) = \|\alpha\|^2$, then

$$
\begin{aligned}
H(t,x,\nabla_x u, \mathrm{Hess}_x u, \alpha) &= b(t,x,\alpha)\cdot\nabla_x u + \frac{1}{2}\mathrm{Tr}(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^T \mathrm{Hess}_x u) + f(t,x,\alpha) \\
&= 2\sqrt{\lambda}\alpha\cdot\nabla_x u + \Delta u + \|\alpha\|^2
\end{aligned}
$$

The HJB equation is given by

$$\partial_t u(t,x) + \min_{\alpha}\left(H(t,x,\nabla_x u, \mathrm{Hess}_x u, \alpha)\right) = 0,$$

where H takes minima at $\alpha = -\sqrt{\lambda}\,\nabla_x u$ and

$$\partial_t u(t,x) + \Delta u - \lambda\|\nabla_x u\|^2 = 0$$

Deep BSDE

Example

Consider a Linear-quadratic Gaussian control Problem $dx_t = 2\sqrt{\lambda}\alpha_t dt + \sqrt{2}dW_t$, with cost functional $J[\alpha] = \mathbb{E}\left[\int_0^T \|\alpha_s\|^2 ds + g(x_T)\right]$.

Here $b(t,x,\alpha) = 2\sqrt{\lambda}\alpha, \sigma(t,x,\alpha) = \sqrt{2}, f(t,x,\alpha) = \|\alpha\|^2$, then

$$
\begin{aligned}
H(t,x,\nabla_x u, \text{Hess}_x u, \alpha) &= b(t,x,\alpha)\cdot\nabla_x u + \frac{1}{2}\text{Tr}(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^T\text{Hess}_x u) + f(t,x,\alpha) \\
&= 2\sqrt{\lambda}\alpha\cdot\nabla_x u + \Delta u + \|\alpha\|^2
\end{aligned}
$$

The HJB equation is given by

$$\partial_t u(t,x) + \min_\alpha \left(H(t,x,\nabla_x u, \text{Hess}_x u, \alpha)\right) = 0,$$

where H takes minima at $\alpha = -\sqrt{\lambda}\,\nabla_x u$ and

$$\partial_t u(t,x) + \Delta u - \lambda\|\nabla_x u\|^2 = 0$$