如何通俗易懂地解释HJB方程?

于66
学精算的科研废物
+ 关注

最开始的回答实在火车上无聊，随手一写，里面一些符号细节也没太在意，以后会慢慢修正，也希望评论区指正。

随着该回答的不断丰富，也希望能随着自己知识的完善不断丰富这个回答，

其实HJB方程很简单，就是一个动态规划一个泰勒展开。

考虑一段时间 $[0,T]$ ，我们的目标是目标函数最大化，通常为一个积分的形式，例如 $\int_0^T U(x(t),u(t),t)dt$ ，

- 其中， $u=\{u(t),t\in[0,T]\}$ 是策略过程，即我们在每个时刻 t 都有一个策略 u 。需要注意一下，我们这里是动态最优化问题，也就是这个决策者在初始时刻对于之后每个时刻都做决策。下述问题除非额外说明，策略过程都属于可行策略集 $u\in\mathcal{A}$ 。（顺便说一下，决策者在初始时对今后每个时刻做决策和决策者在每个时刻都做一个决策是两个不同的策略过程，但是呢，当贝尔曼原则适用时，这两个策略过程是一样的，这也是为什么我们可以应用动态规划来处理这个问题，下文会详细解释）
- $\{x(t),t\in[0,T]\}$ 是状态过程，表示在该时刻的状态。

1. 先来讲动态规划

考虑如下最优化问题

目标函数为 $\max_u \int_t^T f(X^u(s), u(s))ds$，这里的 T 不是随机变量。

初始值 $X(t) = x$

状态过程：$dX^u(t) = g(t, X(t), u(t))dt$

令值函数为 $V(t, x) = \max_u \int_t^T f(X^u(s), u(s))ds$

终值条件：$V(T, X(T)) = 0$

先来说一下动态规划，当贝尔曼原则满足的时候，我们就可以用动态规划来求解问题。其实也就是决策者在任何时刻做的最优决策都是一致的。举个例子：

我作为一个决策者，

- 我在初始时刻 t 做了一个策略 $\{u(s), s \in [t, T]\}$ 这个策略最大化 $[t, T]$ 时间段的目标函数，就是 $\max_U \int_t^T f(X^u(s), u(s))ds$，
- 然后我在 $t+h$ 时刻做了第二组策略 $\{\hat{u}(s), s \in [t+h, T]\}$，这个策略最大化 $[t+h, T]$时间段的目标函数，就是 $\max_U \int_{t+h}^T f(X^u(s), u(s))ds$，

在任何时刻 $m \in [t+h, T]$ 内都满足 $u(m) = \hat{u}(m)$ 时，我们就说这是时间一致的，那么就满足贝尔曼原则，就可以用动态规划了。当然时间不一致的情况也存在，时间一致和不一致取决于目标函数的形式。如果是时间不一致的情况下，我们需要采用其他方法推导HJB。现在假设时间一致，就是满足贝尔曼原则，也就是说可以用动态规划的方式。

根据动态规划我们可以将目标函数写作

$$
\begin{aligned}
V(t,x) &= \max_u \int_t^T f(X^u(s),u(s))ds \\
&= \max_u \{ \int_t^{t+h} f(X^u(s),u(s))ds + \max_u \int_{t+h}^T f(X^u(s),u(s))ds \} \\
&= \max_u \{ \int_t^{t+h} f(X^u(s),u(s))ds + V(t+h, X^u(t+h)) \}
\end{aligned}
$$

这时候我们发现我们可以把 $[t, t+h]$ 时间区间内的控制策略择出来了，他只出现在 $\int_t^{t+h} f(X^u(s),u(s))ds$ 和 $V(t+h, X^u(t+h))$ 中的$X^u(t+h)$ 中。

当 $h \to 0$ 是，这个式子就是 t 时刻的控制 $u(t)$ 的函数，此时我们把问题从求解一组 $\{u(s), s \in [t,T]\}$ 转化为求解 t 时刻的控制 $u(t)$ 了。

注意，$V(t+h, X^u(t+h))$ 是 $t+h$ 之后时间的值函数。它已经是最优的了，和 $\{u(s), s \in [t+h, T]\}$ 无关。只和 $\{u(s), s \in [t, t+h]\}$ 有关，因为这决定了 $X^u(t+h)$ 。

2. 泰勒展开

在得到了下式之后

$$
V(t,x) = \max_u \{ \int_t^{t+h} f(X^u(s),u(s))ds + V(t+h, X^u(t+h)) \}
$$

如果令 $h \to 0$, 要使 $\{\int_t^{t+h} f(X^u(s),u(s))ds + V(t+h, X^u(t+h)\}$ 取得最大值，我们需要对 $u(t)$ 求导，并令导数为0以求得使上式达到最大值时的 $u^*(t)$ 。

但是 $V(t+h, X^u_{t+h})\}$ 对 $u(t)$ 求导不好表示，我们需要利用泰勒展开将其在 $V(t,x)$ 处展开

$$
V(t+h, X^u(t+h)) = V(t,x) + \int_t^{t+h} V_s ds + \int_t^{t+h} V_x dX^u(s) + \frac{1}{2} \int_t^{t+h} V_{xx} (dX^u(s))^2
$$

因为这样做的话，就可以把 $V(t+h, X^u(t+h))$ 用初始值 $V(t,x)$ 和 控制 $\{u(s), s\in[t,t+h]\}$ 表示出来。

3. HJB方程

将 $V(t+h, X^u(t+h))$ 代入 $V(t,x)=\int_t^{t+h} f(X^u(s),u(s))ds+V(t+h,X^u(t+h))$ 中，等式两边同时除以 h 令 $h\to 0$ 就可以得到HJB方程了

$$V(t,x)=\max_u\{\int_t^{t+h} f(X^u(s),u(s))ds+V(t+h,X^u(t+h))\}$$
$$V(t,x)=\max_u\{\int_t^{t+h} f(X^u(s),u(s))ds+V(t,x)+\int_t^{t+h} V_t ds+\int_t^{t+h} V_x dX^u(s)\}$$
$$0=\max_u\{\int_t^{t+h} f(X^u(s),u(s))ds+\int_t^{t+h} V_t ds+\int_t^{t+h} V_x dX^u(s)\}$$
$$0=\max_u\{\int_t^{t+h} f(X^u(s),u(s))+V_t+V_x g(t,X(s),u(s))\}ds$$

根据之前的讨论，我们现在需要做的就是令 $h\to 0$ (目的在于把 $\{u(s), s\in[t,t+h]\}$ 缩成 $u(t)$ 这一个点的策略), 就得到了传说中 HJB方程：

$$0=\max_u\{f(x,u(t))+V_t+V_x g(t,X(t),u(t))\}$$

首先我们 $f(x,u(t))+V_t+V_x g(t,X(t),u(t))$ 对 $u(t)$ 求导，令其为0，求得最优的控制 $u^*(t)$.

然后再把 $u^*(t)$ 代回HJB 方程中， 就变成了 V 的微分方程，再加上终值条件 $V(T,x(T))=0$ 就可以求解得出值函数 $V(t,x)$ 。

我们来观察观察这个HJB方程，他代表的是初始的这个小时刻，在最优的策略下，向后走一点点时间，一方面，会给目标函数一点积累量 $f(x,u(t))$ ，同时，他也会使得今后的值函数产生点变化 V_t , 另一方面，他会影响下一点点时刻的状态，而这个状态的一点点变化，也会影响今后的值函数 $V_x g(t,X(t),u(t))$。所以说，HJB方程就是随着时间的推移，在最优的策略下，他的值函数不变。

回过头来再去思考一下一维的静态最优化问题，静态最优化问题就是一个条件：一阶导为0。

而动态最优化问题的条件变成了两个：（1）在每个时刻一阶导为0。（2）目标从一而终——值函数随着时间的推移不发生变化。

3. 带折现因子怎么办

看到评论中有问带折现因子怎么办，我们只要搞懂问题的本质，其实一通百通，这里来一起算一下。

考虑最优化问题

$\max_{u\in\mathcal{A}}\{\int_t^T e^{-\beta s}f(u(s),X^u(s))ds\}$

控制：$u=\{u(s),s\in[t,T]\},u\in\mathcal{A}$，这里 $\mathcal{A}$ 是可行策略集。

状态过程：$dX^u(t)=g(t,X^u(t),u(t))dt$

值函数：$V(t,x)=\max_{u\in\mathcal{A}}\{\int_t^T e^{-\beta s}f(u(s),X^u(s))ds\}$

边界条件：$V(T,X(T))=0$

动态规划

$$\max_u\{\int_t^{t+h}e^{-\beta t}f(u(t),x(t))+e^{-\beta h}V(t+h,X^u(t+h))\}$$

泰勒展开

$de^{-\beta t}V(t,X^u(t))=e^{-\beta t}[-\beta Vdt+V_tdt+V_xdX^u(t)]$

代进去

$$
\begin{aligned}
V(t,x) &= \max_u \{ \int_t^{t+h} e^{-\beta s} f(u(s),x(s))ds + e^{-\beta(t+h)} V(t+h, X^u(t+h)) \} \\
&= \max_u \{ \int_t^{t+h} e^{-\beta s} f(u(s),x(s))ds + V(t,x) + \\
&\qquad \int_t^{t+h} e^{-\beta s} [-\beta V ds + V_t ds + V_x dX^u(s)]
\end{aligned}
$$

HJB 方程

$\max_u \{ e^{-\beta t} f(u(t),x) - \beta V + V_t + V_x g(t,x,u(t)) \} = 0$

---

4. 如果终止点有额外的效用（就只改变边界条件）

目标函数:

$$
\max_{u \in \mathcal{U}} \int_{t_0}^{T} f[t,x(t),u(t)]dt + G(T,x(T))
$$

$$
\begin{cases} dx(t) = g(t,x(t),u(t))dt \\ x(t_0) = x_0 \end{cases}
$$

定义： $V(t,x) = \max_{u \in \mathcal{U}} \int_t^T f[s,x(s),u(s)]ds + G(T,x(T))$

则有: $V(T,x(T)) = G(T,x(T))$

动态规划

$$
\begin{aligned}
V(t_0,x_0) &= \max_{u(t),t\in[t_0,T]} \int_{t_0}^{T} f[t,x(t),u(t)]dt + G(T,x(T)) \\
&= \max_{u(t),t\in[t_0,T]} \int_{t_0}^{t_0+h} f[t,x(t),u(t)]dt + \int_{t_0+h}^{T} f[t,x(t),u(t)]dt + G(T,X(T)) \\
&= \max_{u(t),t\in[t_0,t_0+h]} \int_{t_0}^{t_0+h} f[t,x(t),u(t)]dt + \max_{u(t),t\in[t_0+h,T]} \int_{t_0+h}^{T} f[t,x(t),u(t)]dt \\
&\qquad + G(T,X(T)) \\
&= \max_{u(t),t\in[t_0,t_0+h]} \int_{t_0}^{t_0+h} f[t,x(t),u(t)]dt + V(t_0+h,x(t_0+h))
\end{aligned}
$$

HJB 方程

$$
\begin{cases}
0 = \max_{u(t)}\{f[t,x(t),u(t)] + V_t(t,x(t)) + V_x(t,x(t))g(t,x(t),u(t))\} \\
V(T,x(T)) = G(T,x(T))
\end{cases}
$$

---

5. 如果状态过程中有随机项（以布朗运动为例）

目标函数：$\max_u \mathbb{E}_{t,x} \int_t^T f(X^u(s),u(s))ds$

状态过程：$dX(t) = g(t,X^u(t),u(t))dt + \sigma dW(t)$

动态规划：

$$
\begin{aligned}
&\max_u \mathbb{E}_{t,x} \int_t^T f(X(s),u(s))ds \\
=&\max_u \mathbb{E}_{t,x}\{\int_t^{t+h} f(X(s),u(s))ds + V(t+h,X(t+h))\}
\end{aligned}
$$

泰勒展开：

$$
\begin{aligned}
dV(t+h, X(t+h)) &= V_t dt + V_x dX(t) + \frac{1}{2} V_{xx} [dX(t)]^2 \\
&= V_t dt + V_x dX(t) + \frac{1}{2} V_{xx} \sigma^2 dt
\end{aligned}
$$

代进去，得到HJB方程：

$$0 = \max_u \{ f(x, u(t)) + V_t + V_x g(t, X(t), u(t)) + \tfrac{1}{2} V_{xx} \sigma^2 \}$$

---

6. 如果终止时间是个随机变量

这个分成两方面来讲，一种是知道T的分布，另一种是T是首达时。

(1)

编辑于 2022-11-28 14:51

曲奇饼
low latency trader

@于66 的回答已经很清楚地推导了HJB方程在连续情况下是什么样子的了，珠玉在前，我补充几个例子。实际上这两个例子展示了HJB方程的动机，也就是它到底是怎么来的。

第一个例子是离散情况下的Bellman Equation，考虑一个时间离散的系统 $t \in [0, 1, \ldots T]$, 每个时间点上有一个状态 $x \in \mathcal{X}$, 同时我们可以在该时刻, 在该状态上进行一个操作 $a \in \mathcal{A}$.

其中 $\mathcal{X}, \mathcal{A}$ 为可能的状态以及操作集合. 操作 a 作用到了状态 x 以后，将转变为一个新的状态 $\hat{x}$, 也就是有状态转移函数:

$f : \mathcal{X} \times \mathcal{A} \to \mathcal{X}, \hat{x} = f(x, a)$

对应得到的回报为:

$r : \mathcal{X} \times \mathcal{A} \to \mathbb{R}, r = r(x, a)$

于是，当我们在每个时间 t 做出了一个操作 π_t ,形成了一个策略 $\pi \in \Pi$ 后，得到的累计回报为：

$R(x_0, \pi) = \sum_{i=0}^{T} r(x_i, \pi_i)$

其中 Π 为所有可能的策略集合. 我们的优化目标就是找到一个最优操作组合 $\pi \in \Pi$, 使得我们的总回报达到最大:

$$\begin{aligned} V(x_0) &= \max_{\pi} R(x_0, \pi) \\ s.t. \quad x_{t+1} &= f(x_t, \pi_t) \end{aligned}$$

这种优化问题我们称为动态规划. 求解动态规划的方法是从结尾 $t = T$ 开始倒推过程。

我们发现，考虑从结束时间往前数 τ 个时间，这段时间的 $t \in [T - \tau, T]$ 也对应了一个动态规划:

$$\begin{aligned} V(\tau, x_{T-\tau}) &= \max_{\pi} R(x_{T-\tau}, \pi_{T-\tau}) \\ &= \max_{\pi} \sum_{i=T-\tau}^{T} r(x_i, \pi_i) \end{aligned}$$

当 $t = T$ 时，这段时间的总回报为

$V(0, x_T) = \max_{\pi} R(x_T, \pi_T) = r(x_T, \pi_T)$

是确定的，换句话说就是边界条件，那只要我们能得到 $V(\tau - 1, x_{T-(\tau-1)})$ 和 $V(\tau, x_{T-\tau})$ 之间的迭代关系，就可以求解出我们想要的问题 $V(x_0) := V(\tau = T, x_0)$ ，这个迭代关系就是 Bellman Equation：

$$
\begin{aligned}
V(\tau, x_{T-\tau}) &= \max_{\pi} R(x_{T-\tau}, \pi_{T-\tau}) \\
&= \max_{\pi} \sum_{i=T-\tau}^{T} r(x_i, \pi_i) \\
&= \max_{\pi} \left\{ r(x_{T-\tau}, \pi_{T-\tau}) + \sum_{i=T-(\tau-1)}^{T} r(x_i, \pi_i) \right\} \\
&= \max_{\pi} \left\{ r(x_{T-\tau}, \pi_{T-\tau}) + \max_{\pi} \sum_{i=T-(\tau-1)}^{T} r(x_i, \pi_i) \right\} \\
&= \max_{\pi} \{ r(x_{T-\tau}, \pi_{T-\tau}) + V(\tau-1, x_{T-(\tau-1)}) \} \\
&= \max_{\pi} \{ r(x_{T-\tau}, \pi_{T-\tau}) + V(\tau-1, f(x_{T-\tau})) \}
\end{aligned}
$$

这种离散形式的动态规划问题例子数不胜数，我们举一个属于leet code简单难度动态规划题目(实际上任意离散系统的动态规划最优算法题目都是这样的解法)：给定一个二叉树，每个Node上有一个score：

```
class TreeNode:
    def __init__(self, val=0, left=None, right=None):
        self.val = val
        self.left = left
        self.right = right
```

求一个从根节点到叶子节点最大score的路径. 设平衡二叉树的深度为 T ，根节点深度为0。$x \in \mathcal{X}$ 为所有的树节点. 所有操作就是从该节点走向left or right. Bellman Equation为

$$
V(\tau, x_{T-\tau}) = \max_{\pi \in [\text{left, right}]} \{ r(x_{T-\tau}, \pi_{T-\tau}) + V(\tau-1, f(x_{T-\tau})) \}
$$

其中状态转移函数以及收益函数为:

$$
\begin{aligned}
f(x, \text{left}) &= \text{left}(x) \\
f(x, \text{right}) &= \text{right}(x) \\
r(x, a) &= \text{value}(x)
\end{aligned}
$$

于是Bellman Equation带入，就得到了以下的solution:

```
def maximum_path(node: TreeNode) -> int:
    if node == None:
        return 0
    return node.score + max(maximum_path(node.left), maximum_path(node.right))
```

对于一个input:

```
      1
    /   \
   0     7
  / \   / \
 1   5 6   5
```

得到的最大score为14, 也就是 root->right->left->terminated.

第二个例子是连续时间系统下，怎么从经典物理学中，由最小作用量原理导出的Hamilton-Jabobi Equation，推广到动态规划问题的Hamilton-Jabobi-Bellman Equation的。

我们回顾一下经典分析力学，这里只说一个框架，详情可以参考朗道的理论物理学教程第一卷。考虑一个经典物理系统的粒子的作用量为:

$$S(q,t) = \int_0^T L(q,\dot{q},t)dt$$

其中 q 是广义坐标， $\dot{q} := dq/dt$ 为广义速度。最小作用量原理告诉我们，一个粒子的运动轨迹 q ,会使得系统的作用量取到最小, 也就是变分原理 $\delta S = 0$,于是得到了拉格朗日方程:

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}}) - \frac{\partial L}{\partial q} = 0$$

其中 $p = \frac{\partial L}{\partial \dot{q}}$ 为广义坐标对应的广义动量， $\frac{\partial L}{\partial q}$ 为广义力。所以这是牛顿第二定律 $F = dp/dt$ 的高级版本。

我们定义哈密顿量为

$H(p,q,t) = p\dot{q} - L(q,\dot{q},t)$

可以得到哈密顿量满足正则方程:

$\dot{q} = \frac{\partial H}{\partial p}, \dot{p} = -\frac{\partial H}{\partial q}$

考虑作用量的全微分:

$$\begin{aligned} dS = Ldt &= \frac{\partial S}{\partial q}dq + \frac{\partial S}{\partial t}dt \\ &= \frac{\partial S}{\partial q}\dot{q}dt + \frac{\partial S}{\partial t}dt \\ &= (p\dot{q} + \frac{\partial S}{\partial t})dt \\ &\Rightarrow \frac{\partial S}{\partial t} + p\dot{q} - L = 0 \end{aligned}$$

注意到哈密顿量的定义，于是便得到了Hamilton-Jacobi方程:

$\frac{\partial S}{\partial t} = -H(p,q,t)$

和动态规划理论里的HJB方程几乎就是一个形式(以Halmilton-Jacobi-Bellman命名，就是因为这是力学系统的推广)。因为我们只要将最小作用量原理考虑为一个最优控制问题

$$\begin{aligned} S = \min_q \int_0^T L(q,\dot{q},t)dt \\ s.t. \quad \dot{q} = f(q,\dot{q},t) \end{aligned}$$

代入HJB方程就可以得到

$0 = \frac{\partial S}{\partial t} + \min_q \left\{ L(q,\dot{q},t) + \frac{\partial S}{\partial q}\dot{q} \right\}$

如果我们取 $H = \min_q \left\{ L(q,\dot{q},t) + \frac{\partial S}{\partial q}\dot{q} \right\}$ ，就得到了哈密顿-雅可比方程 $\frac{\partial S}{\partial t} + H = 0$

具体怎么将Hamilton-Jacobi Equation推广到HJB Equation，就很自然了。

在力学中，我们通过操作速度，控制曲线的运动轨迹，使得系统的作用量达到最小。而在一个一般化的情况：

$$\begin{aligned} \text{作用量} &\longrightarrow \text{收益}(\max) \text{ or } \text{成本}(\min) \\ \text{坐标} &\longrightarrow \text{状态 } x \\ \text{速度} &\longrightarrow \text{操作 } \pi \end{aligned}$$

这里用累计收益为例，我们定义Bellman函数为累计收益的最大值：

$$\begin{aligned} V(x,t) = \max_{\pi} R(x_0,\pi) = g(x(T)) + \int_t^T r(x(s),\pi(s))ds \\ s.t. \quad \dot{x} = f(x,\pi) \end{aligned}$$

其中 R 的第一项是终止收益，第二项的积分是时间累计收益，于是回报函数 $r(x,\pi)$ 就是一个广义的拉格朗日量，累计收益就是广义的作用量，于是我们定义广义的动量为

$p = \frac{\partial V}{\partial x}$

定义广义的哈密顿量

$H(p,x,t) = f(x^*,\pi^*)\cdot p^* + r(x^*,\pi^*)$,

同样可以得到广义的正则方程：

$\dot{x}^* = \frac{\partial H}{\partial p}, \dot{p}^* = -\frac{\partial H}{\partial x}$

其中*号表示最优控制下的路径。可以证明，Bellman函数 $V(x,t)$ 满足以下的PDE

$$\begin{cases} \dfrac{\partial V}{\partial t} + H(x,p,t) = 0 \\ V(x,T) = g(x) \end{cases}$$

这就是HJB方程，这个结果称为Pontryagin Maximum Principle。更多的细节可以参考J. Yong 和 X. Y. Zhou 写的Stochastic Controls Hamiltonian Systems and HJB Equations

进一步地，我们将动态规划原理，从确定性的体系可以推广到不确定的体系。在具有随机性的体系中，我们寻求一个最优控制 π^* ,使得目标收益的期望值取到最大:

$$
\begin{aligned}
& V(x,t) = \max_{\pi} \mathbb{E}_{t,x}\left[G(X_t^{\pi}) + \int_t^T F(X_s^{\pi},\pi_s,s)ds\right] \\
s.t. \quad & dX_t^{\pi} = \mu(X_t^{\pi},\pi_t,t)dt + \sigma(X_t^{\pi},\pi_t,t)dW_t
\end{aligned}
$$

其中 X_t 为伊藤过程, W_t 为布朗运动。 这里就不能直接套用确定性系统里的玩意了，因为得用伊藤引理来做微分：

$$
\begin{aligned}
dV &= \left(\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}\mu + \frac{1}{2}\frac{\partial^2 V}{\partial x^2}\sigma^2\right)dt + \frac{\partial V}{\partial x}\sigma dW_t \\
V(x,t) &= \max_{\pi} \mathbb{E}_{x,t}\left[V(x_{t+dt},t+dt) + F(X_t^{\pi},\pi_t,t)dt\right] \\
&\Rightarrow \frac{\partial V}{\partial t} + \max_{\pi}\left\{\frac{\partial V}{\partial x}\mu + \frac{1}{2}\frac{\partial^2 V}{\partial x^2}\sigma^2 + F(X_t^{\pi},\pi_t,t)\right\} = 0
\end{aligned}
$$

第一行是伊藤引理，第二行是无穷小形式的动态规划原理，第三行就是最终得到的HJB方程。伊藤引理中的不确定项在求平均值的时候被消掉了：

$$\mathbb{E}_{x,t}\left\{dW_t\right\} = 0$$

但观察HJB方程的形式，我们仍然可以定义哈密顿量

$$H(x,t) = \max_{\pi}\left\{\frac{\partial V}{\partial x}\mu + \frac{1}{2}\frac{\partial^2 V}{\partial x^2}\sigma^2 + F(X_t^{\pi},\pi_t,t)\right\}$$

得到与Hamilton-Jacobi方程完全一样的形式。

编辑于 2023-06-24 20:54

Kagami今晚吃什么
抱歉，不感兴趣

+ 关注

7 人赞同了该回答

随手写一个回答吧，看到这个HJB了吗

$$V_t + \sup_u (V_x + F) = 0$$

这个式子表示如果过去了很小的一个时间间隔dt，那么你有一个收入F，其次还有V在这个dt里的变化，首先是 V_t ，其次状态变量X的变化也与t有关，因此还有 V_x （如果X是个随机过程，这里还需要加个用Ito formula可以推出来的 V_{xx} ）。

那么这个式子的经济学含义是啥，或者说如何通俗理解这个式子呢?

假设你现在有个矿山，价值是V，V和时间t和状态变量X有关（这个状态变量我们不妨设成环境咯）。现在我们有一种开矿策略u（control variable），这种开矿策略会连续地给我们带来一笔收入，即F。同时这种开矿策略也会影响到环境X，从而间接影响到V。

那么HJB的意思就是，如果我们采取了最优的开矿策略u，那么我们在t这个时点获得的收入F，以及矿山价值的变化（ $V_t + V_x$ ），加起来应该是0。

为什么？因为这本质上就是在左口袋倒右口袋（矿山在左口袋，矿山产生的收入进了右口袋），你是不可能获得正收益的，即你左口袋加上右口袋的资产的加总应该是不变的。所谓的最优策略只是想办法让这个事情无损耗的发生而已。

当然如果我们考虑贴现（即资金的时间价值）的情况的话，就会得到在文献中更常见的那个HJB了，即

$$V_t + \sup_u (V_x + F) = rV$$

这个经济学含义是，你左右口袋的资产的增长率只能是无风险收益率。

发布于 2023-06-20 19:54

Naenia Dea
谢绝不可接触者巡逻，谢谢配合。

12 人赞同了该回答

我不知道所谓“通俗易懂”的解释是什么意思，如果你是想问HJB方程的本质是什么，我觉得本质是Itō drift-diffusion process上的动态规划。因为HJB方程就是用价值函数推出来的。如果你问的是推导过程，那翻翻书大概都有；自己定义一个价值函数来推的话，只要记住“随机过程不能用泰勒展开而要使用伊藤引理”多试几次基本都能推出来。（我本来想放一个推导过程在这里但是你乎公式编辑器犯病了我嗯是写不出公式。）

问题就在于这种已经不算基础的数学知识还要求通俗易懂的解释，这怎么可能做得到啊......

发布于 2020-09-09 10:02

▲ 赞同 12 ▼ 1 条评论 分享 收藏 喜欢 ···

Mr Figurant
中央财经大学 金融学博士在读
+ 关注

3 人赞同了该回答

HJB方程（Hamilton-Jacobi-Bellman，汉密尔顿-雅可比-贝尔曼方程）是在扩展种类模型（Expanding variety models）中引入的，即由发明者决定在研发上投入多少：

$$\dot{V}(v,t) = r(t)V(v,t) - \pi(v,t)$$

HJB简单推导如下，先从 $t+\Delta t$ 拆分：

$$\begin{aligned} V(v,t) = &\int_t^{t+\Delta t} \exp\left(-\int_t^s r(\tau)d\tau\right)[p^x(v,s)-\psi]x(v,s)ds \\ &+\int_{t+\Delta t}^{\infty} \exp\left(-\int_t^{t+\Delta t} r(\tau)d\tau - \int_{t+\Delta t}^s r(\tau)d\tau\right)[p^x(v,s)-\psi]x(v,s)ds \end{aligned}$$

对于微小的 $\Delta t, s \in [t, t+\Delta t]$ ，对于连续函数 f 有 $f(s) \approx f(t)$ ，因此：

> $V(v,t) = [px(v,t) - \psi]x(v,t)\Delta t + \exp(-r(t)\Delta t)V(v,t+\Delta t) + o_1(\Delta t)$

得到：

> $[px(v,t) - \psi]x(v,t)\Delta t + (1 - r(t)\Delta t)V(v,t+\Delta t) - V(v,t) + o_2(\Delta t) = 0$

两端同时除以 Δt ，令 $\Delta t \to 0$ ，若 $V(s,\cdot)$ 可微，则得到 $\dot{V}(v,t) = r(t)V(v,t) - \pi(v,t)$ 。

学习更多：

Mr Figurant：高级宏观07：扩展种类模型
2 赞同 · 2 评论　文章

发布于 2023-10-03 10:12