Clase 4.2 Aproximación de la función de Valor

- Deep Mind
- GitHub
- Implementar
  - de Ft. Valor, ∇

} Ayudantía

Una forma de aproximar funciones de Valor (Regresión lineal)

$$y = \theta^T \phi(x)$$

$$[\ \ldots\ ]\begin{bmatrix} \vdots \\ \vdots \\ \vdots \end{bmatrix}$$

$$y(x) = \theta_0 + \theta_1 x + \theta_2 x_2 + \ldots + \theta_p x_p$$

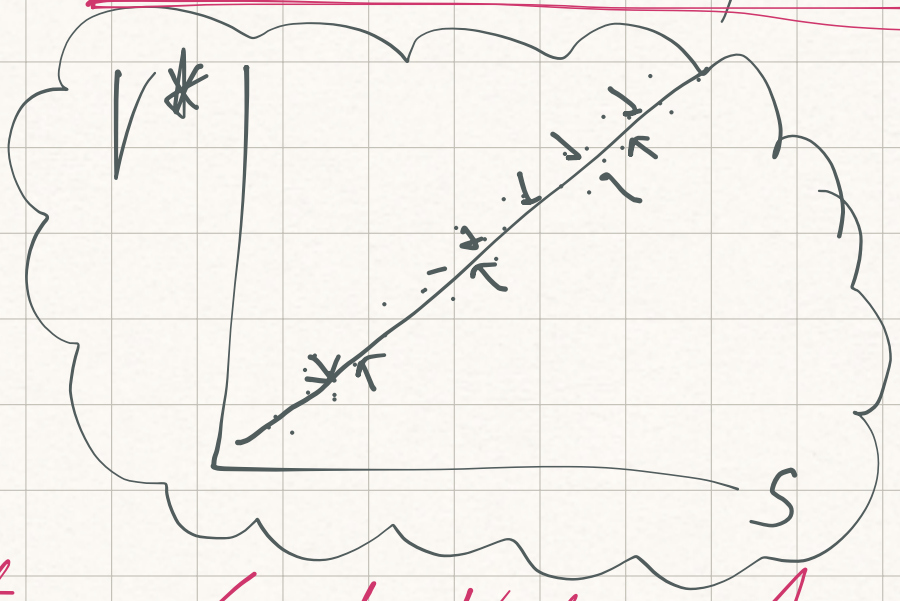
$$\phi(x) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

$$\phi(x) = \begin{bmatrix} x \\ x\theta \\ x^2\theta \\ \vdots \end{bmatrix}$$

(features) → atributos

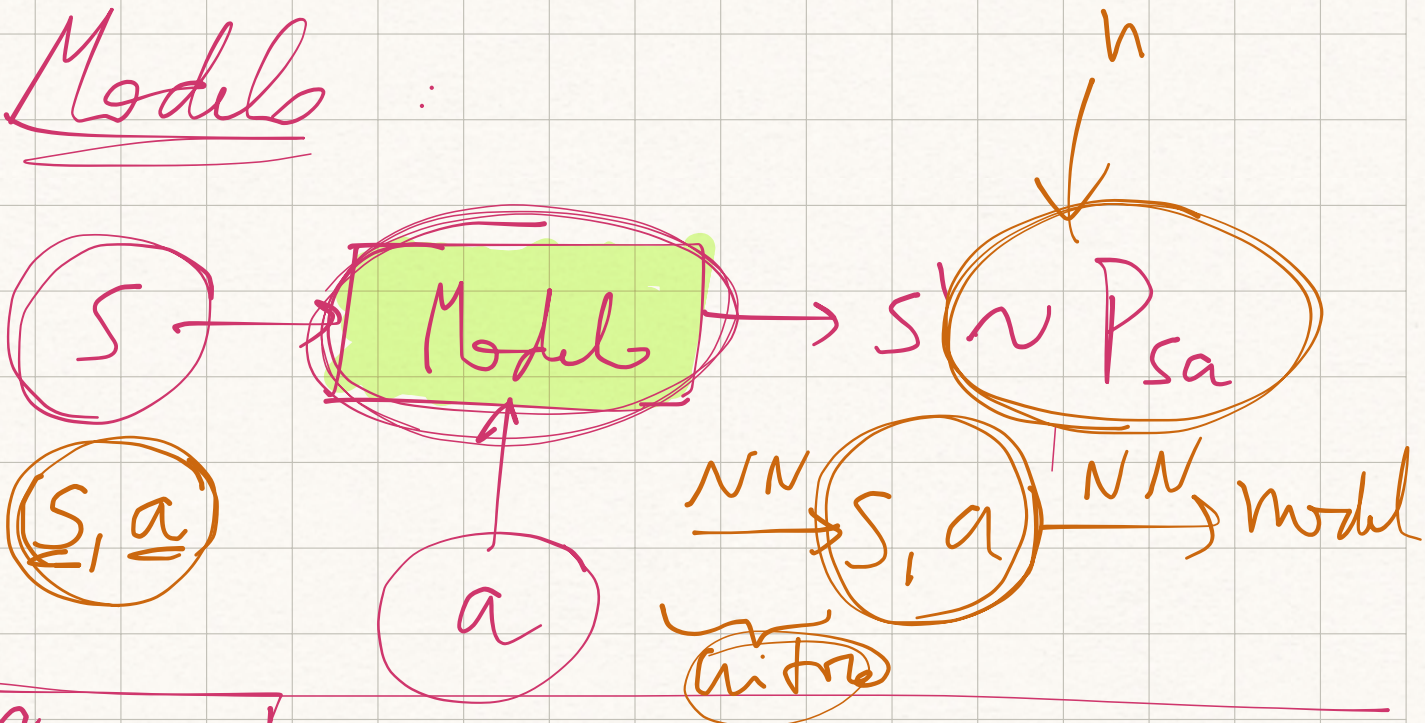
Aproximación de fn Valor, en est. continuos

$$V^*(s) \simeq \theta^T \phi(s)$$



Iteración de Valor Ajustada (IVA)

1) Modelo :



* Asume que el conjunto de accion

(A) es discreto, pequeño.

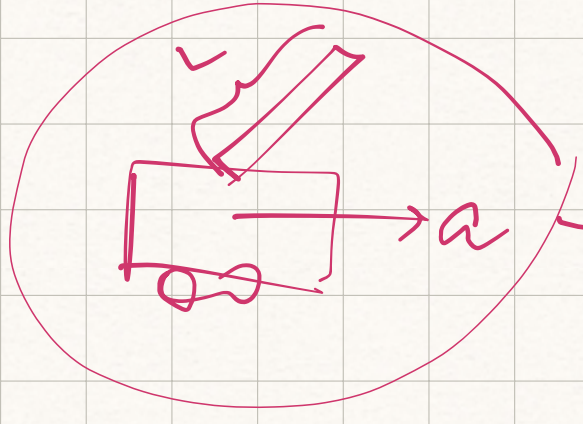
en gral, $\dim A < \dim S$

	A	S
Conjunto	2	4
AA	2	6
Pelo	4	12

---

i) Simulador físico (dinámico)

$$S = \begin{pmatrix} x \\ \dot{x} \\ \theta \\ \dot{\theta} \end{pmatrix} \rightarrow \boxed{S' = \begin{pmatrix} \dot{x} \\ \alpha - L\beta\cos(\theta)/m \\ \dot{\theta} \\ \beta \end{pmatrix}}$$



Determinista (tener o no tener miedo)

ii) Aprendiendo desde la Data

$$s_0^{(1)} \xrightarrow{a_0} s_1^{(1)} \xrightarrow{a_1^{(1)}} s_2^{(1)} \xrightarrow{a_2^{(1)}} \cdots \rightarrow s_T^{(1)}$$

$$S^{(m)} \xrightarrow{a_0^{(m)}} S_1^{(m)} \xrightarrow{a_1^{(m)}} S_2 \xrightarrow{a_2^{(m)}} \cdots \rightarrow S_T^{(m)}$$

$$S_{t+1} = S, a$$

$$\boxed{S_{t+1} = A S_t + B a_t}$$

→ fn. lineal determinística

→

$$\boxed{S' = f(S, a) + \varepsilon}$$

ε: ruido

Prob. trans. fn lineal estocástica

Data set

$$\{[(S_0^{(1)}, a_0^{(1)}), S_1^{(1)}] \quad \ldots\ldots \text{m veces hasta} \quad [(S_{T-1}^{(1)}, a_{T-1}^{(1)}), S_T^{(1)}]$$

$(S_0^{(1)}, a_0^{(1)})$: x ; $S_1^{(1)}$: y ; $(S_{T-1}^{(1)}, a_{T-1}^{(1)})$: x ; $S_T^{(1)}$: y

Regresión lineal

$$\min_{A,B} \sum_{i=0}^{n} \sum_{T=0}^{T-1} \left\| S_{T-1}^{(i)} - A_{ST}^{(i)} - B_{a_T}^{(i)} \right\|_2$$

Enfoques para funciones no lineales

$$f_\theta \text{ (parametrizado x } \theta)$$

$$f_\theta(s,a) = \underline{A}\,\phi_1(s) + \underline{B}\,\phi_2(a) + \varepsilon$$

atrib. Estado

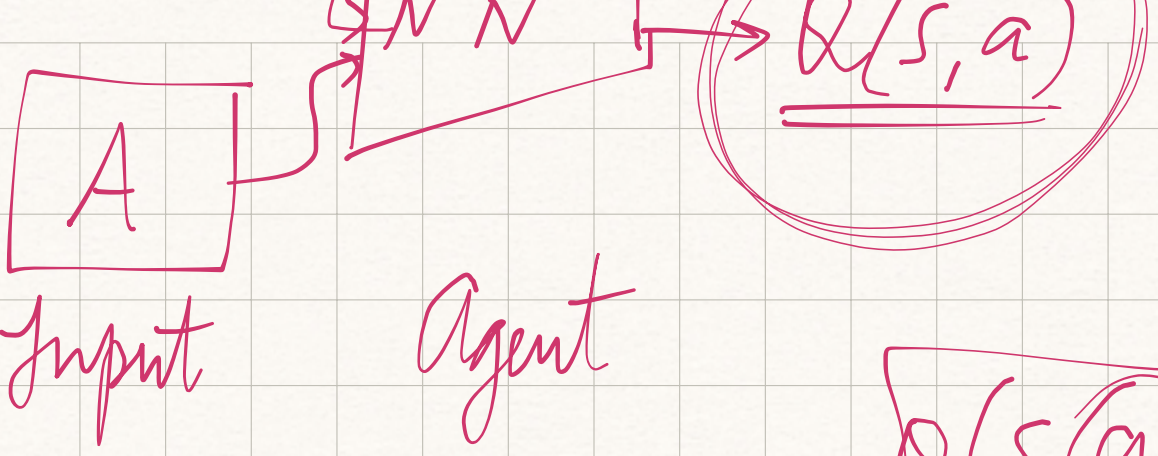
atrib. acción

Deep Reinf. Learning (DRL)

$$f_\theta(s,a) = nn(s,a)$$

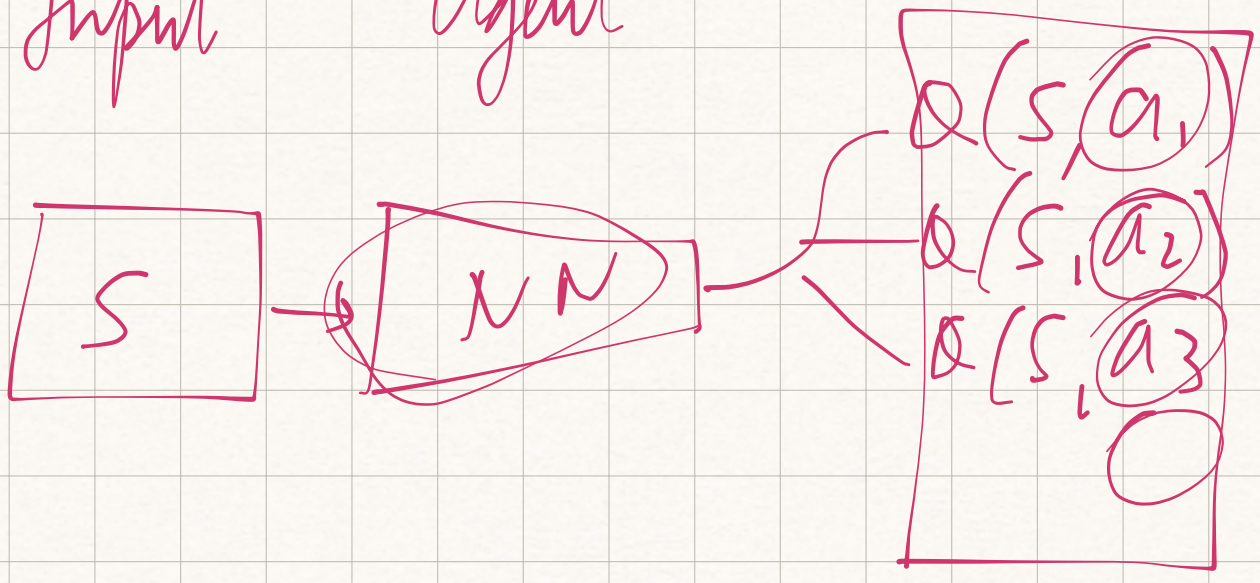
Redes Neuronales

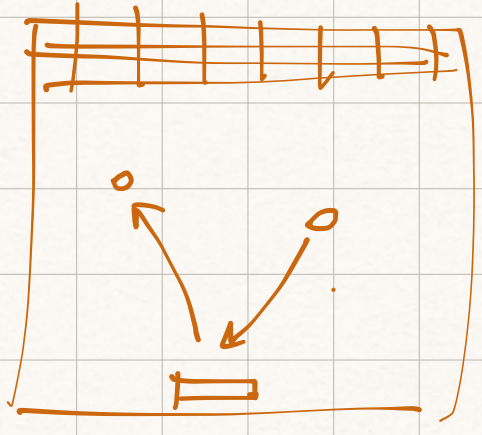
i) [s]

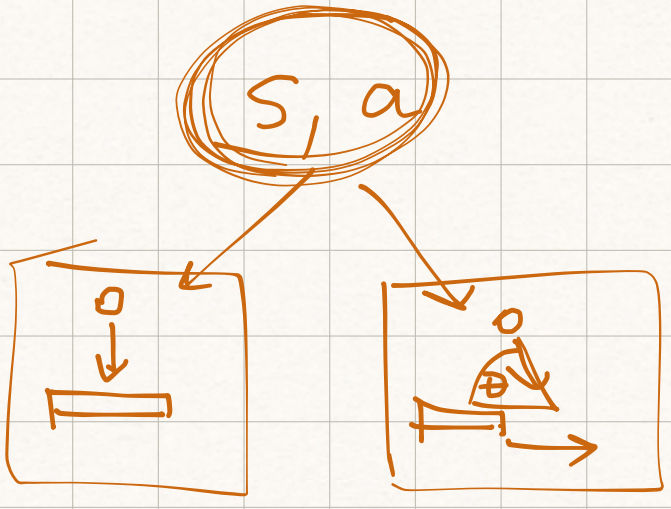


Input Agent

ii)







$$\phi_1(s) = \begin{bmatrix} x \\ x\theta \\ x^2 \\ \theta \\ x\theta \end{bmatrix}$$