Lecture 7
Function approximation

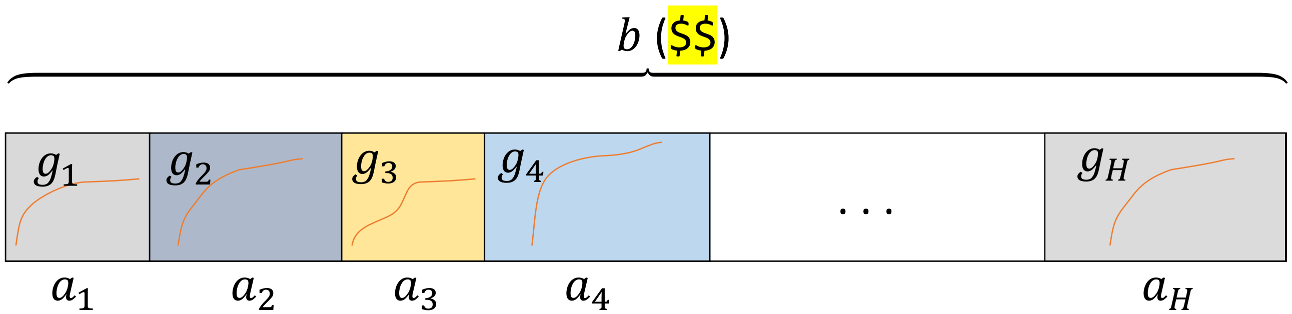
Year 1963

Polynomial Approximation—A New Computational Technique in Dynamic Programming: Allocation Processes

By Richard Bellman, Robert Kalaba, and Bella Kotkin

Mathematics and Computation 17:155—161, 1963

The problem & the dynamic programming formulation



$$v_1^*(0) = \max\{\, \Sigma_{i=1}^{H}\, g_i(a_i) : a_1, \ldots a_H \geq 0, a_1 + \cdots + a_H = b\} =?$$

$a_1^*, \ldots, a_H^* =?$

$v_h^*(s)$: optimal value achievable over $[h, H]$
when $s \in [0, b]$ of the resource is used beforehand

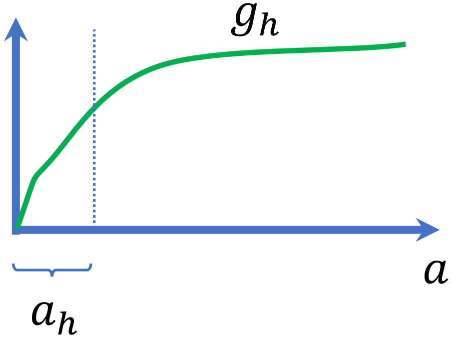
$$v_{H+1}^*(s) = 0$$

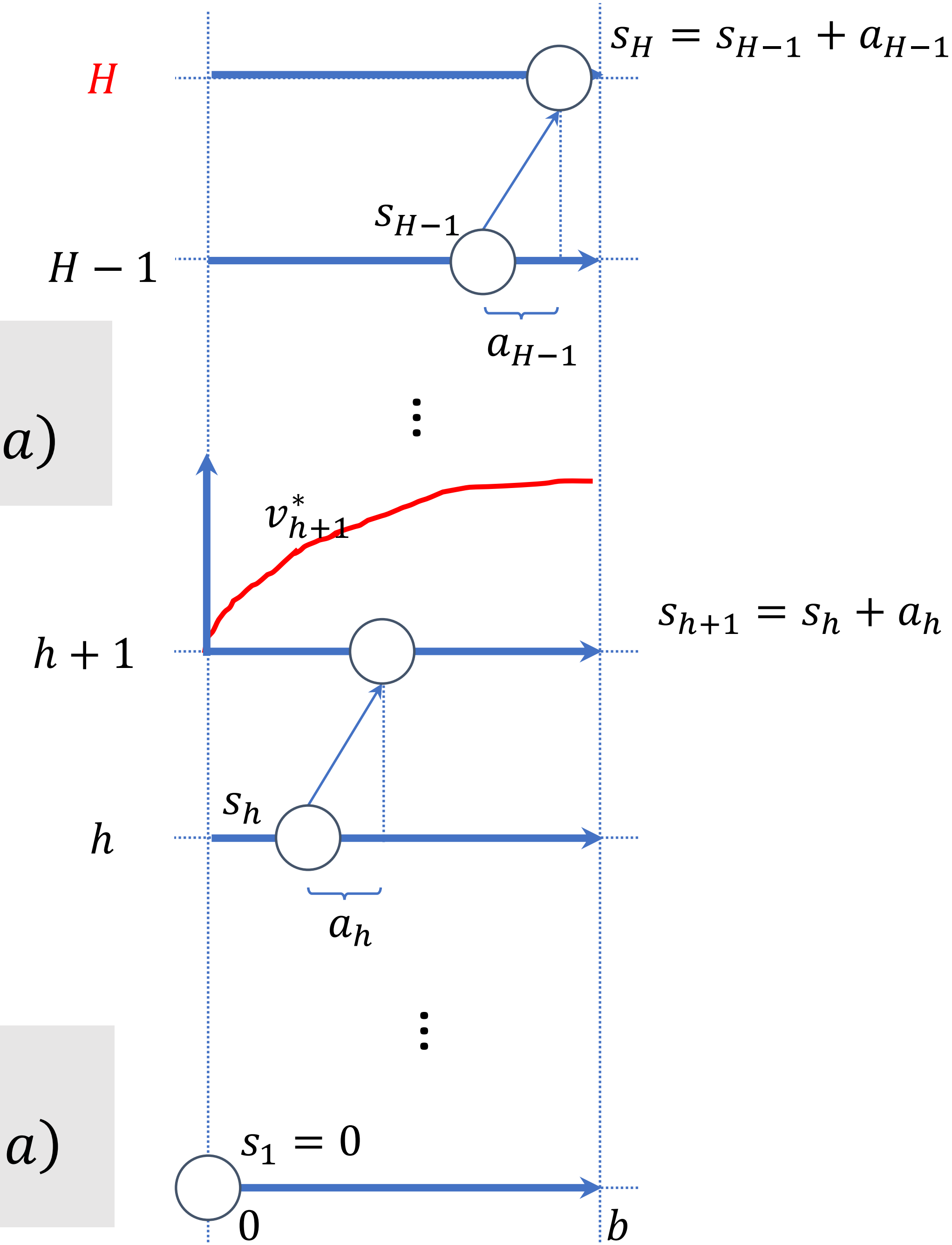
$$v_h^*(s) = \max_{0 \leq a \leq b-s} \; g_h(a) + v_{h+1}^*(s+a) \qquad\qquad 1 \leq h \leq H$$

$$v^*_{H+1}(s) = 0$$

for $h = H, H-1, \dots, 1$

$$v^*_h(s) = \max_{0 \le a \le b-s} g_h(a) + v^*_{h+1}(s+a)$$





$$\pi^*_h(s) = \operatorname*{argmax}_{0 \le a \le b-s} g_h(a) + v^*_{h+1}(s+a)$$

Hov

$v_h^*(s$ ... $\leq h$

$\leq H$

$s \in \{0$

on

what if $a_1, \ldots, a_H \in [0, b]^p$?

need $\left(\frac{b}{\epsilon}\right)^p$ storage, compute, brr…

"curse of dimensionality"

https://www.rand.org/content/dam/rand/pubs/reports/2006/R352.pdf

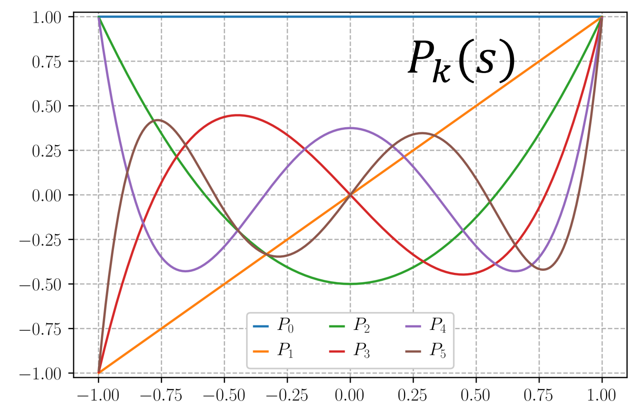
New idea (in 1963): Generalized polynomial approximation

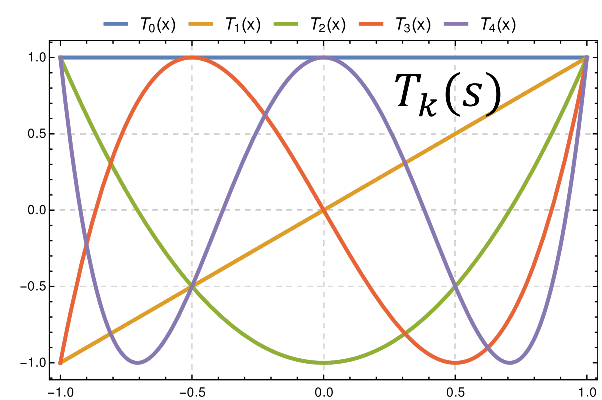
$$f(x) = \Sigma_{k=1}^{d} \theta_k \varphi_k(s), \qquad s \in [-1,1]$$

If using $\{P_k\}$, orthonormal set w.r.t. uniform measure on [-1,1]

$$\theta_k = \int_{-1}^{1} f \varphi_k$$

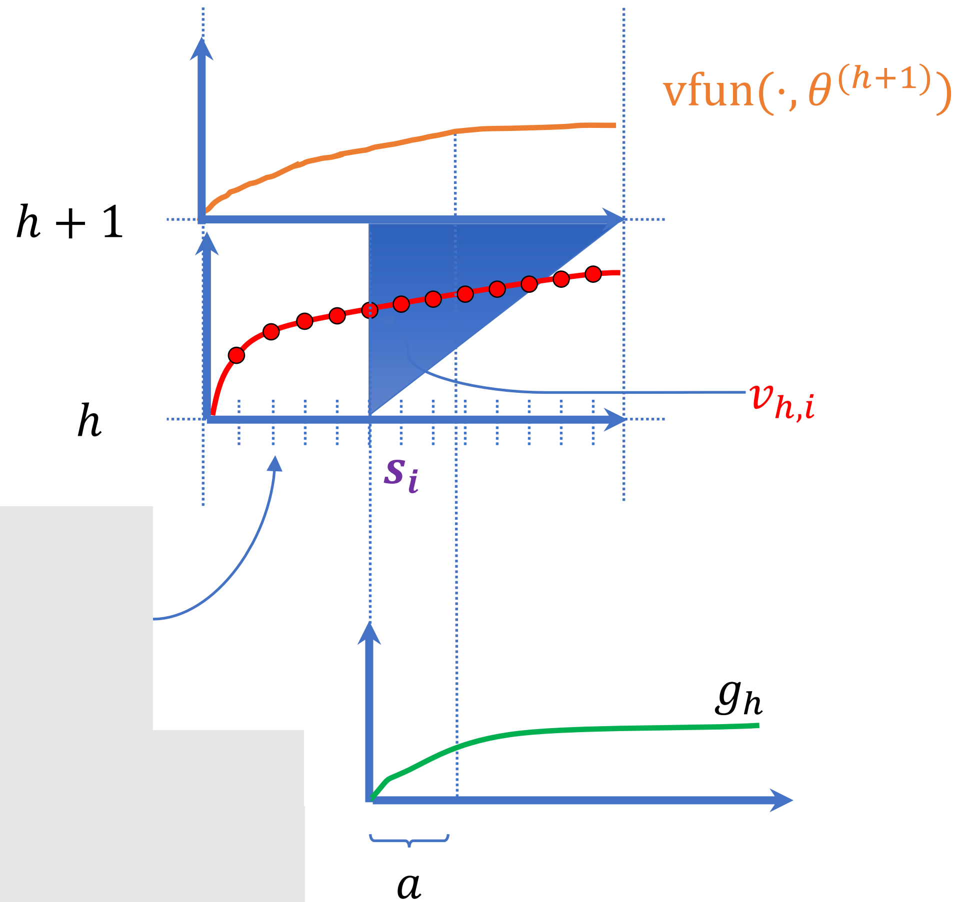
+Gauss quadratures: eval f at $\{s_i\}_{i\in[r]}$





https://en.wikipedia.org/wiki/Legendre_polynomials

https://en.wikipedia.org/wiki/Chebyshev_polynomials



precompute $\boldsymbol{s_1}, \ldots, \boldsymbol{s_r} \in [0, b]^p$

and $w_1, \ldots, w_r \in \mathbb{R}$

$\theta^{(H+1)} := (0, \ldots, 0) \in \mathbb{R}^d$

for $h = H, H-1, \ldots, 1$:

  for $i \in [r]$:

    $v_{h,i} := \max_{0 \le a \le b - \boldsymbol{s_i}} g_h(a) + \text{vfun}\left(\boldsymbol{s_i} + a, \theta^{(h+1)}\right)$

  for $j \in [d]$:

    $\theta_j^{(h)} := \Sigma_{i=1}^{r} w_i v_{h,i} \varphi_j(\boldsymbol{s_i})$

$$\text{vfun}(s, \theta) := \Sigma_{k=1}^{d} \theta_k \, \varphi_k(s)$$

FVI
= Fitted value iteration

~DQN

What did we gain?

Storage: $O(H(d+r))$
Compute cost: $O(H(d+\ r\ \mathrm{OptCost})\ )$

Compare with $\left(\frac{b}{\epsilon}\right)^{p}$!

Gain?

For fixed OptCost, the cost is independent of the dimension p

At least we can run the procedure:
Fine-grained error control through the choice of $\{\phi_k\}_k, r$ and OptCost

Relevance?

$$v_h^*(s) = \max_{0 \le a \le b-s} g_h(a) + v_{h+1}^*(s+a)$$

$$\Rightarrow$$

$$v_h^*(s) = \max_{a \in \mathcal{A}_h(s)} r_h(s,a) + \mathbb{E}_\xi[v_{h+1}^*(f_h(s,a,\xi))]$$

Markov Decision Processes

The RL Hypothesis

Dynamic programming +
function approximation

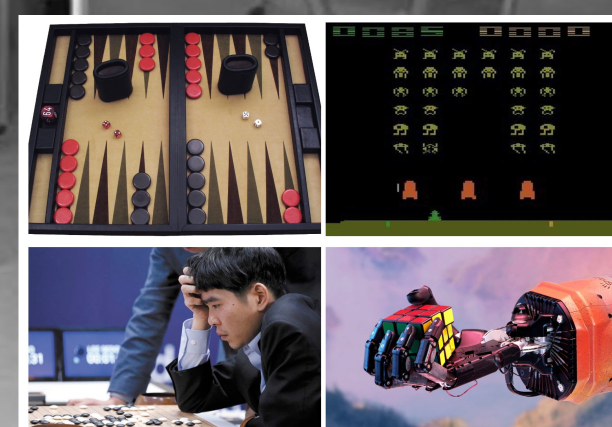
=

key technique to solve large scale
~~control problems~~

Markov Decision Problems

Empirical tests

“Finally, if we combine these techniques – polynomial approximations and Lagrange multipliers – with that of successive approximations, there should be very few allocation processes which still resist our efforts.”



IBM 7090 “supercomputer”

But.. does this work..?

1. Approximation: How large should be the degree of polynomials used to approximate v^*?

   Smoothness, approximation theory, systems theory..

2. Computation:
   Given that we can approximate well v^*, say,

   $v^*(s) = \Sigma_{i=1}^{d} \theta_i^* \varphi_i(s)$,

   how much computation is needed to get $\theta^* = (\theta_1^*, \ldots, \theta_d^*)$?

   Can we do it in $\mathrm{poly}(A, H, d, 1/\varepsilon)$ regardless of dimension (state space size)?

O' Curse of Dimensionality, Where is Thy Sting?

Kenneth L. Judd
Hoover Institution and NBER

April 11, 2008

https://kenjudd.org/wp-content/uploads/2017/02/Curse_in_Dallas.pdf

Math Tool II: Efficient Function Approximation

- Linear polynomial methods:

$$f(x, y, z, ...) = \sum_{i=1}^{m} a_i \phi_i (x, y, z, ...)\ ,\ \phi_i \text{ multivariate polynomials}$$

  – Simple tensor product approach produces approximations like

$$\sum_{i=0}^{m} \sum_{j=0}^{m} \sum_{k=0}^{m} a_{ij} x^i y^j z^k$$

  – Proper notion of "degree" in multivariate context is sum of powers

$$\text{degree}\left(x^i y^j z^k\right) = i + j + k$$

  – Complete polynomials like

$$\sum_{i+j+k \leq m} a_{ijk} x^i y^j z^k$$

  have far fewer terms by a ratio of nearly $d!$, but are almost as good

  – See Gaspar-Judd (1997)

https://kenjudd.org/wp-content/uploads/2016/09/jgasweb.pdf
https://www.cambridge.org/core/journals/macroeconomic-dynamics/article/abs/solving-largescale-rationalexpectations-models/A484F77266454AA52B535BF8B28257B8

Modeling assumptions

- $v^* \in \mathcal{F}$ realizable
- $q^* \in \mathcal{F}'$realizable
- $v^\pi \in \mathcal{F}$ for any deterministic/stochastic ML π
- $q^\pi \in \mathcal{F}'$ for any deterministic/stochastic ML π
- $T^\pi \mathcal{F} \subset \mathcal{F}, T^\pi \mathcal{F}' \subset \mathcal{F}'$ for any deterministic/stochastic ML π
- $T\mathcal{F} \subset \mathcal{F}, T\mathcal{F}' \subset \mathcal{F}'$
- $T^\pi \mathbb{R}^{\mathcal{S}} \subset \mathcal{F}, T^\pi \mathbb{R}^{\mathcal{S} \times A} \subset \mathcal{F}'$ for any deterministic/stochastic ML π
- $T\mathbb{R}^{\mathcal{S}} \subset \mathcal{F}, T\mathbb{R}^{\mathcal{S} \times A} \subset \mathcal{F}'$

Questions from slack

Matthew Pietrosanu 19 hours ago

I'll bite since this relates to some of my research in functional data analysis. The choice of basis {phi_j: j=1,...,d} isn't discussed much in the notes. Are there any particularly common choices in RL? (I can only speak to statistics, so I'm curious.) Or am I reading too much into what will ultimately be just a "toy" model (e.g., with the bases obtained by other means, say, NN?)

+9

Matthew Pietrosanu 19 hours ago

As a followup, though this involves infinite-dimensional state spaces.. Are there any settings where estimating this basis is a primary concern? (e.g., a smooth basis that describes some "optimal" d-dimensional subset of the infinite-dimensional space of functions.) In such a setting, estimating Phi (as a matrix) along some grid in S might not be adequate (e.g., the basis described by Phi may be non-smooth). Are there any approaches in RL to deal with this? (Again, maybe this isn't even a relevant problem.)

+2