Artificial Intelligence and Machine Learning (AIML)

2023–24

- **This lecture**: extended DP lecture to ensure solid grasp of concepts (change of plan)
- **Next lecture**: logic

Dynamic programming: extended lecture

- **Three levels of sophistication**
  a) **Apply** Bellman recurrence to a particular problem, and/or identify known Bellman recurrence
  b) **Derive** Bellman recurrence from SDP requires (even if implicit/intuitive) understanding relationship between SDP and recurrence
  c) **Rigorous understanding** of the *principle of optimality*
- **In this course** (a) **competence** on simple examples, (b) **optional** (will show examples today), (c) **optional** – ask if interested.

Dynamic programming: extended lecture

- **Level (a):** Apply Bellman recurrence

(Bellman recursion, specific case, LN 5.6)

$$x = [2, -3, -4, 6, 5, -1]$$

$$P^*_0 = 0$$

$$P^*_n = \max(0, P^*_{n-1} + x_n)$$

$$X^*_0 = [\ ]$$

$$X^*_n = [\ ] \text{ if } P^*_{n-1} + x_n \leq 0,\ X^*_n \cup [x_n] \text{ otherwise}$$

(Bellman recursion, specific case, LN 5.6)

$x = [2, -3, -4, 6, 5, -1]$

$P^*_0 = 0$

$P^*_n = \max(0, P^*_{n-1} + x_n)$

$X^*_0 = [\,]$

$X^*_n = [\,]$ if $P^*_{n-1} + x_n \leq 0$, $X^*_n \cup [x_n]$ otherwise

$n=0$: $P^*_0 = 0$, $X^*_0 = [\,]$

(Bellman recursion, specific case, LN 5.6)

$x = [2, -3, -4, 6, 5, -1]$

$P^*_0 = 0$

$P^*_n = \max(0, P^*_{n-1} + x_n)$

$X^*_0 = [\,]$

$X^*_n = [\,]$ if $P^*_{n-1} + x_n \leq 0$, $X^*_n \cup [x_n]$ otherwise

n=0: $P^*_0 = 0$, $X^*_0 = [\,]$

n=1: $P^*_1 = \max(0, P^*_0 + x_1) = \max(0, 2) = 2$, $X^*_1 = [2]$

(Bellman recursion, specific case, LN 5.6)

$x = [2, -3, -4, 6, 5, -1]$

$P^*_0 = 0$

$P^*_n = \max(0, P^*_{n-1} + x_n)$

$X^*_0 = [\ ]$

$X^*_n = [\ ]$ if $P^*_{n-1} + x_n \leq 0$, $X^*_n \cup [x_n]$ otherwise

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

n=1: $P^*_1 = \max(0, P^*_0 + x_1) = \max(0, 2) = 2$, $X^*_1 = [2]$

n=2: $P^*_2 = \max(0, P^*_1 + x_2) = \max(0, -1) = 0$, $X^*_2 = [\ ]$

(Bellman recursion, specific case, LN 5.6)

$x = [2, -3, -4, 6, 5, -1]$

$P^*_0 = 0$

$P^*_n = \max(0, P^*_{n-1} + x_n)$

$X^*_0 = [\ ]$

$X^*_n = [\ ]$ if $P^*_{n-1} + x_n \leq 0$, $X^*_n \cup [x_n]$ otherwise

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

n=1: $P^*_1 = \max(0, P^*_0 + x_1) = \max(0, 2) = 2$, $X^*_1 = [2]$

n=2: $P^*_2 = \max(0, P^*_1 + x_2) = \max(0, -1) = 0$, $X^*_2 = [\ ]$

n=3: $P^*_3 = \max(0, P^*_2 + x_3) = \max(0, -4) = 0$, $X^*_3 = [\ ]$

(Bellman recursion, specific case, LN 5.6)

$x = [2, -3, -4, 6, 5, -1]$

$P^*_0 = 0$

$P^*_n = \max(0, P^*_{n-1} + x_n)$

$X^*_0 = [\ ]$

$X^*_n = [\ ]$ if $P^*_{n-1} + x_n \leq 0$, $X^*_n \cup [x_n]$ otherwise

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

n=1: $P^*_1 = \max(0, P^*_0 + x_1) = \max(0, 2) = 2$, $X^*_1 = [2]$

n=2: $P^*_2 = \max(0, P^*_1 + x_2) = \max(0, -1) = 0$, $X^*_2 = [\ ]$

n=3: $P^*_3 = \max(0, P^*_2 + x_3) = \max(0, -4) = 0$, $X^*_3 = [\ ]$

n=4: $P^*_4 = \max(0, P^*_3 + x_4) = \max(0, 6) = 6$, $X^*_4 = [6]$

(Bellman recursion, specific case, LN 5.6)

$x = [2, -3, -4, 6, 5, -1]$

$P^*_0 = 0$

$P^*_n = \max(0, P^*_{n-1} + x_n)$

$X^*_0 = [\ ]$

$X^*_n = [\ ]$ if $P^*_{n-1} + x_n \leq 0$, $X^*_n \cup [x_n]$ otherwise

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

n=1: $P^*_1 = \max(0, P^*_0 + x_1) = \max(0, 2) = 2$, $X^*_1 = [2]$

n=2: $P^*_2 = \max(0, P^*_1 + x_2) = \max(0, -1) = 0$, $X^*_2 = [\ ]$

n=3: $P^*_3 = \max(0, P^*_2 + x_3) = \max(0, -4) = 0$, $X^*_3 = [\ ]$

n=4: $P^*_4 = \max(0, P^*_3 + x_4) = \max(0, 6) = 6$, $X^*_4 = [6]$

n=5: $P^*_5 = \max(0, P^*_4 + x_5) = \max(0, 11) = 11$, $X^*_5 = [6, 5]$

(Bellman recursion, specific case, LN 5.6)

$x = [2, -3, -4, 6, 5, -1]$

$P^*_0 = 0$

$P^*_n = \max(0, P^*_{n-1} + x_n)$

$X^*_0 = [\,]$

$X^*_n = [\,]$ if $P^*_{n-1} + x_n \leq 0$, $X^*_n \cup [x_n]$ otherwise

n=0: $P^*_0 = 0$, $X^*_0 = [\,]$

n=1: $P^*_1 = \max(0, P^*_0 + x_1) = \max(0, 2) = 2$, $X^*_1 = [2]$

n=2: $P^*_2 = \max(0, P^*_1 + x_2) = \max(0, -1) = 0$, $X^*_2 = [\,]$

n=3: $P^*_3 = \max(0, P^*_2 + x_3) = \max(0, -4) = 0$, $X^*_3 = [\,]$

n=4: $P^*_4 = \max(0, P^*_3 + x_4) = \max(0, 6) = 6$, $X^*_4 = [6]$

n=5: $P^*_5 = \max(0, P^*_4 + x_5) = \max(0, 11) = 11$, $X^*_5 = [6, 5]$

n=6: $P^*_6 = \max(0, P^*_5 + x_6) = \max(0, 10) = 10$, $X^*_6 = [6, 5, -1]$

(Bellman recursion, specific case, Tutorial 5 Q1)

$\Pr = [2, 4, 7, 3, 9]$

$P^*_0 = 0$

$P^*_n = \max_{i \in \{1, 2, ..., n\}}(P^*_{n-i} + \Pr_i)$

$X^*_0 = [\,]$

$X^*_n = X^*_{n-i} \cup [i]$ if $P^*_{n-i} + \Pr_i \geq P^*_{n-j} + \Pr_j$ for all $j \in \{1, 2, ..., n\}$

(Bellman recursion, specific case, Tutorial 5 Q1)

$\mathrm{Pr} = [2, 4, 7, 3, 9]$

$P^*_0 = 0$

$P^*_n = \max_{i \in \{1, 2, ..., n\}}(P^*_{n-i} + \mathrm{Pr}_i)$

$X^*_0 = [\ ]$

$X^*_n = X^*_{n-i} \cup [i]$ if $P^*_{n-i} + \mathrm{Pr}_i \geq P^*_{n-j} + \mathrm{Pr}_j$ for all $j \in \{1, 2, ..., n\}$

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

(Bellman recursion, specific case, Tutorial 5 Q1)

$\mathrm{Pr} = [2, 4, 7, 3, 9]$

$P^*_0 = 0$

$P^*_n = \max_{i \in \{1, 2, ..., n\}}(P^*_{n-i} + \mathrm{Pr}_i)$

$X^*_0 = [\ ]$

$X^*_n = X^*_{n-i} \cup [i]$ if $P^*_{n-i} + \mathrm{Pr}_i \geq P^*_{n-j} + \mathrm{Pr}_j$ for all $j \in \{1, 2, ..., n\}$

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

n=1: $P^*_1 = \max(P^*_0 + \mathrm{Pr}_1) = \max(2) = 2$, $X^*_1 = [1]$

(Bellman recursion, specific case, Tutorial 5 Q1)

$\mathrm{Pr} = [2, 4, 7, 3, 9]$

$P^*_0 = 0$

$P^*_n = \max_{i \in \{1, 2, ..., n\}}(P^*_{n-i} + \mathrm{Pr}_i)$

$X^*_0 = [\ ]$

$X^*_n = X^*_{n-i} \cup [i]$ if $P^*_{n-i} + \mathrm{Pr}_i \geq P^*_{n-j} + \mathrm{Pr}_j$ for all $j \in \{1, 2, ..., n\}$

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

n=1: $P^*_1 = \max(P^*_0 + \mathrm{Pr}_1) = \max(2) = 2$, $X^*_1 = [1]$

n=2: $P^*_2 = \max(P^*_1 + \mathrm{Pr}_1, P^*_0 + \mathrm{Pr}_2) = \max(4, 2) = 4$, $X^*_2 = [2]$

(Bellman recursion, specific case, Tutorial 5 Q1)

$\mathrm{Pr} = [2, 4, 7, 3, 9]$

$P^*_0 = 0$

$P^*_n = \max_{i \in \{1, 2, ..., n\}}(P^*_{n-i} + \mathrm{Pr}_i)$

$X^*_0 = [\ ]$

$X^*_n = X^*_{n-i} \cup [i]$ if $P^*_{n-i} + \mathrm{Pr}_i \geq P^*_{n-j} + \mathrm{Pr}_j$ for all $j \in \{1, 2, ..., n\}$

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

n=1: $P^*_1 = \max(P^*_0 + \mathrm{Pr}_1) = \max(2) = 2$, $X^*_1 = [1]$

n=2: $P^*_2 = \max(P^*_1 + \mathrm{Pr}_1, P^*_0 + \mathrm{Pr}_2) = \max(4, 2) = 4$, $X^*_2 = [2]$

n=3: $P^*_3 = \max(P^*_2 + \mathrm{Pr}_1, P^*_1 + \mathrm{Pr}_2, P^*_0 + \mathrm{Pr}_3) = \max(6, 6, 7) = 7$, $X^*_3 = [3]$

(Bellman recursion, specific case, Tutorial 5 Q1)

$\mathrm{Pr} = [2, 4, 7, 3, 9]$

$P^*_0 = 0$

$P^*_n = \max_{i \in \{1, 2, ..., n\}}(P^*_{n-i} + \mathrm{Pr}_i)$

$X^*_0 = [\ ]$

$X^*_n = X^*_{n-i} \cup [i]$ if $P^*_{n-i} + \mathrm{Pr}_i \geq P^*_{n-j} + \mathrm{Pr}_j$ for all $j \in \{1, 2, ..., n\}$

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

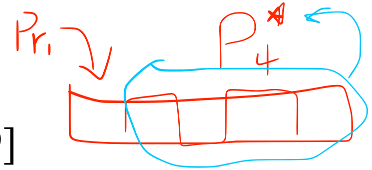
n=1: $P^*_1 = \max(P^*_0+\mathrm{Pr}_1) = \max(2) = 2$, $X^*_1 = [1]$

n=2: $P^*_2 = \max(P^*_1+\mathrm{Pr}_1, P^*_0+\mathrm{Pr}_2) = \max(4, 2) = 4$, $X^*_2 = [2]$

n=3: $P^*_3 = \max(P^*_2+\mathrm{Pr}_1, P^*_1+\mathrm{Pr}_2, P^*_0+\mathrm{Pr}_3) = \max(6, 6, 7) = 7$,
$X^*_3 = [3]$

n=4: $P^*_4 = \max(P^*_3+\mathrm{Pr}_1, P^*_2+\mathrm{Pr}_2, P^*_1+\mathrm{Pr}_3, P^*_0+\mathrm{Pr}_4)$
$= \max(9, 8, 9, 3) = 9$, $X^*_4 = [1,3]$

(Bellman recursion, specific case, Tutorial 5 Q1)



Pr: Price

$\Pr = [2, 4, 7, 3, 9]$

$P^*_0 = 0$ P: Profit P_{n-i} ^ * is the optimal solution for the rest of configurations

$P^*_n = \max_{i \in \{1, 2, ..., n\}} (P^*_{n-i} + \Pr_i)$

X: configuration of cutting units

$X^*_0 = [\ ]$

$X^*_n = X^*_{n-i} \cup [i]$ if $P^*_{n-i} + \Pr_i \geq P^*_{n-j} + \Pr_j$ for all $j \in \{1, 2, ..., n\}$

X_n ^ * This represents the maximum revenue achievable by cutting a plant with a total length of n units.

for example X_1 ^ * = [1, 3] means cut units 1 and units 3.

The condition P_{n-i} ^ * + Pr_i >= P_{n-j} ^ * + Pr_j in the Bellman Recurrence for your plant cutting problem ensures you're always selecting the cut that leads to the maximum total revenue

By considering all possible cuts (j), you compare the sub-problems (P_(n-j)) and ensure you choose the cut that leads to the overall optimal solution (X_n).

n=0: $P^*_0 = 0$, $X^*_0 = [\ ]$

n=1: $P^*_1 = \max(P^*_0 + \Pr_1) = \max(2) = 2$, $X^*_1 = [1]$?

n=2: $P^*_2 = \max(P^*_1 + \Pr_1, P^*_0 + \Pr_2) = \max(4, 2) = 4$, $X^*_2 = [2]$

n=3: $P^*_3 = \max(P^*_2 + \Pr_1, P^*_1 + \Pr_2, P^*_0 + \Pr_3) = \max(6, 6, 7) = 7$, $X^*_3 = [3]$

n=4: $P^*_4 = \max(P^*_3 + \Pr_1, P^*_2 + \Pr_2, P^*_1 + \Pr_3, P^*_0 + \Pr_4)$
$= \max(9, 8, 9, 3) = 9$, $X^*_4 = [1,3]$

n=5: $P^*_5 = \max(P^*_4 + \Pr_1, P^*_3 + \Pr_2, P^*_2 + \Pr_3, P^*_1 + \Pr_4, P^*_0 + \Pr_5)$
$= \max(11, 11, 11, 5, 9) = 11$, $X^*_5 = [1,3,1]$?

Dynamic programming: extended lecture

- **Level (b):** Bellman recurrence from SDP

(Bellman recursion, general form LN 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in S_n} F\left(X'\right)$$

(Bellman recursion, general form LN 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in S_n} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P.)

$$S_n = \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}$$

(Bellman recursion, general form LN 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in S_n} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P.)

$$S_n = \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}$$

(Insert into 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}} F\left(X'\right)$$

(Bellman recursion, general form LN 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in S_n} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P.)

$$S_n = \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}$$

(Insert into 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}} F\left(X'\right)$$

(Expand min range)

$$= \min\left(F\left([\,]\right), F\left(X_{n-1}^\star \cup [x_n]\right)\right)$$

(Bellman recursion, general form LN 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in S_n} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P.)

$$S_n = \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}$$

(Insert into 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}} F\left(X'\right)$$

(Expand min range)

$$= \min\left(F\left([\,]\right), F\left(X_{n-1}^\star \cup [x_n]\right)\right)$$

(Apply obj. function)

$$= \min\left(0, F\left(X_{n-1}^\star\right) + x_n\right)$$

(Bellman recursion, general form LN 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in S_n} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P.)

$$S_n = \left\{[\,]\,, X_{n-1}^\star \cup [x_n]\right\}$$

(Insert into 5.3)

$$F\left(X_n^\star\right) = \min_{X' \in \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}} F\left(X'\right)$$

(Expand min range)

$$= \min\left(F\left([\,]\right), F\left(X_{n-1}^\star \cup [x_n]\right)\right)$$

(Apply obj. function)

$$= \min\left(0, F\left(X_{n-1}^\star\right) + x_n\right)$$

(Change of notation)

$$P_n^\star = F\left(X_n^\star\right) \qquad P_n^\star = \min\left(0, P_{n-1}^\star + x_n\right)$$

(Bellman recursion, general form LN 5.3) $$F\left(X_n^\star\right) = \min_{X' \in S_n} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P.) $$S_n = \left\{[\,],X_{n-1}^\star \cup [x_n]\right\}$$

(Insert into 5.3) $$F\left(X_n^\star\right) = \min_{X' \in \left\{[\,],X_{n-1}^\star \cup [x_n]\right\}} F\left(X'\right)$$

(Expand min range) $$= \min\left(F\left([\,]\right), F\left(X_{n-1}^\star \cup [x_n]\right)\right)$$

(Apply obj. function) $$= \min\left(0, F\left(X_{n-1}^\star\right) + x_n\right)$$

(Change of notation) $$P_n^\star = F\left(X_n^\star\right) \qquad P_n^\star = \min\left(0, P_{n-1}^\star + x_n\right)$$

(SDP initialization) $$X_0^\star = [\,] \qquad P_0^\star = 0$$

(Bellman recursion, general form LN 5.4)

$$X_n^\star = \operatorname*{argmin}_{X' \in S_n} F\left(X'\right)$$

AIML

(Bellman recursion, general form LN 5.4)

$$X_n^{\star} = \underset{X' \in S_n}{\operatorname{argmin}} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P., insert into 5.4)

$$X_n^{\star} = \underset{X' \in \left\{[\,], X_{n-1}^{\star} \cup [x_n]\right\}}{\operatorname{argmin}} F\left(X'\right)$$

(Bellman recursion, general form LN 5.4)

$$X_n^\star = \operatorname*{argmin}_{X' \in S_n} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P., insert into 5.4)

$$X_n^\star = \operatorname*{argmin}_{X' \in \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}} F\left(X'\right)$$

(Expand min range)

$$= \operatorname{argmin}\left(F\left([\,]\right), F\left(X_{n-1}^\star \cup [x_n]\right)\right)$$

(Bellman recursion, general form LN 5.4)

$$X_n^\star = \underset{X' \in S_n}{\operatorname{argmin}} F(X')$$

(Tail subsequence SDP update, P.O.P., insert into 5.4)

$$X_n^\star = \underset{X' \in \{[\,], X_{n-1}^\star \cup [x_n]\}}{\operatorname{argmin}} F(X')$$

(Expand min range)

$$= \operatorname{argmin}\left(F([\,]), F\left(X_{n-1}^\star \cup [x_n]\right)\right)$$

(Cases/conditional)

$$= \begin{cases} [\,] & F([\,]) < F\left(X_{n-1}^\star \cup [x_n]\right) \\ X_{n-1}^\star \cup [x_n] & \text{otherwise} \end{cases}$$

(Bellman recursion, general form LN 5.4)

$$X_n^\star = \underset{X' \in S_n}{\operatorname{argmin}} F\left(X'\right)$$

(Tail subsequence SDP update, P.O.P., insert into 5.4)

$$X_n^\star = \underset{X' \in \left\{[\,], X_{n-1}^\star \cup [x_n]\right\}}{\operatorname{argmin}} F\left(X'\right)$$

(Expand min range)

$$= \operatorname{argmin}\left(F\left([\,]\right), F\left(X_{n-1}^\star \cup [x_n]\right)\right)$$

(Cases/conditional)

$$= \begin{cases} [\,] & F\left([\,]\right) < F\left(X_{n-1}^\star \cup [x_n]\right) \\ X_{n-1}^\star \cup [x_n] & \text{otherwise} \end{cases}$$

(Apply obj. function, change of notation)

$$= \begin{cases} [\,] & 0 < P_{n-1}^\star + x_n \\ X_{n-1}^\star \cup [x_n] & \text{otherwise.} \end{cases}$$

(Bellman recursion, general form LN 5.4)

$$X_n^\star = \underset{X' \in S_n}{\operatorname{argmin}} F(X')$$

(Tail subsequence SDP update, P.O.P., insert into 5.4)

$$X_n^\star = \underset{X' \in \{[], X_{n-1}^\star \cup [x_n]\}}{\operatorname{argmin}} F(X')$$

(Expand min range)

$$= \operatorname{argmin}\left(F([]), F\left(X_{n-1}^\star \cup [x_n]\right)\right)$$

(Cases/conditional)

$$= \begin{cases} [] & F([]) < F\left(X_{n-1}^\star \cup [x_n]\right) \\ X_{n-1}^\star \cup [x_n] & \text{otherwise} \end{cases}$$

(Apply obj. function, change of notation)

$$= \begin{cases} [] & 0 < P_{n-1}^\star + x_n \\ X_{n-1}^\star \cup [x_n] & \text{otherwise.} \end{cases}$$

(SDP initialization)

$$X_0^\star = []$$

Dynamic programming: extended lecture

- **FAQ1**: Since DP is *exact*, is it actually *exhaustive*?
- **A**: No: principle of optimality (PoP) means DP Bellman recursion *follows the same recursive structure* as the exhaustive SDP but computing optimal values, using only a *single sequence of optimal solutions* in each stage

Dynamic programming: extended lecture

- **FAQ2**: If DP only retains *one optimal configuration* at each stage, why is it *not greedy*?
- **A**: Greedy algorithms must compute configurations at each stage, whereas DP does not need to compute configurations themselves (due to PoP), DP can re-use optimal solutions from earlier stages (*memoization*)

Dynamic programming: extended lecture

- **FAQ3**: *Why* does DP work?
- **A**: PoP with SDP: objective function is a *homomorphism* from SDP to optimal values, *preserving distributivity* in the SDP

AIML

Dynamic programming: extended lecture

- **FAQ4**: What is *memoization*, exactly?
- **A**: Retaining optimal configurations at previous stages, *not essential* for DP (essential: PoP with SDP)

Dynamic programming: extended lecture

- **FAQ5**: What about *overlapping subproblems*?
- **A**: Multiple configurations in later stages make use of same solution obtained earlier, often cited to distinguish DP from *divide-and-conquer* but *not essential* (essential: PoP with SDP)