Lecture 5. Optimal Control

Deterministic Infinite Time

Controller Design Problem (Infinite Horizon)

- Given a state space **X**, a control space **U**, an observation space **Y**, an initial state, noise models in stochastic case
- We need to design a control law that minimizes the objective function

	Deterministic Model	**Stochastic Model**
Continuous time	$\dot{x} = f(x,u)$ $y = h(x,u)$ $J = \int_{t=0}^{\infty} \delta^t l_t(x,u,\dot{x})dt$	$\dot{x} = f(x,u,w)$ $y = h(x,u,v)$ $J = \int_{t=0}^{\infty} \delta^t l_t(x,u,\dot{x})dt$
Discrete time	$x_{k+1} = f(x_k,u_k)$ $y_k = h(x_k,u_k)$ $J = \sum_{k=0}^{\infty} \delta^k l_k(x_k,u_k,x_{k+1})$	$x_{k+1} = f(x_k,u_k,w_k)$ $y_k = h(x_k,u_k,v_k)$ $J = \sum_{k=0}^{\infty} \delta^k l_k(x_k,u_k,x_{k+1})$

Infinite Time Optimal Control

- Stationarity
- Solving stationary Bellman equation
- Linear quadratic regulator

Infinite Horizon vs Finite Horizon

- Two main features of infinite horizon problems:
  - The number of time steps is infinite.
  - The system is stationary, i.e., the system equation and the run-time cost does not change from time to time.

Infinite Horizon vs Finite Horizon

- The optimal policies under infinite horizon are often simpler than their finite horizon counterparts.

- Infinite horizon problems generally require a more sophisticated mathematical treatment.

Problem Formulation

Discrete Time

$$J = \sum_{k=0}^{\infty} \delta^k l(x_k, u_k)$$

$$x_{k+1} = f(x_k, u_k)$$

Continuous Time

$$J = \int_{t=0}^{\infty} \delta^t l(x, u) dt$$

$$\dot{x} = f(x, u)$$

$$V(k, x_k) = \min_{u_k, u_{k+1} \ldots} \sum_{k'=k}^{\infty} \delta^{k'-k} l(x_{k'}, u_{k'})$$

$$V(t, x_t) = \min_{u_{t:\infty}} \int_{\tau=t}^{\infty} \delta^{\tau-t} l(x, u) d\tau$$

The value function should be time-invariant

Stationarity

- Since the dynamics and the run-time cost is stationary, starting from an initial state x^* at any step k will yield the same optimal control sequence, hence the same value function.

$$
\begin{aligned}
V(k, x^*) &= \min_{u_k, u_{k+1}\dots} \{l(x^*, u_k) + \sum_{k'=k+1}^{\infty} \delta^{k'-k} l(x_{k'}, u_{k'})\} \\
&= \min_{u_0, u_1\dots} \{l(x^*, u_0) + \sum_{k'=1}^{\infty} \delta^{k'} l(x_{k'}, u_{k'})\} \\
&= V(0, x^*)
\end{aligned}
$$

- The value function does not depend on time (in both discrete time and in continuous time). In the following discussion, we will drop the time dependency in the value function.

Bellman Equation in Infinite Horizon

- Definition of value function

$$
\begin{aligned}
V(x_k) &= \min_{u_k, u_{k+1}\dots} \sum_{k'=k}^{\infty} \delta^{k'-k} l(x_{k'}, u_{k'}) \\
&= \min_{u_k, u_{k+1}\dots} \{l(x_k, u_k) + \sum_{k'=k+1}^{\infty} \delta^{k'-k} l(x_{k'}, u_{k'})\} \\
&= \min_{u_k} \{l(x_k, u_k) + \delta \min_{u_{k+1},\dots} \sum_{k'=k+1}^{\infty} \delta^{k'-k-1} l(x_{k'}, u_{k'})\}
\end{aligned}
$$

- Bellman equation with discount

$$
V(x_k) = \min_{u_k} \{l(x_k, u_k) + \delta \cdot V(x_{k+1})\}
$$

HJB Equation in Infinite Horizon

- Definition of value function

$$\begin{aligned} V(x_t) &= \min_{u_{t:\infty}} \int_{\tau=t}^{\infty} \delta^{\tau-t} l(x,u) d\tau \\ &= \min_{u_{t:t'}} \{ \int_{\tau=t}^{t'} \delta^{\tau-t} l(x,u) d\tau + \delta^{t'-t} V(x_{t'}) \} \end{aligned}$$

- Taking the limit

$$0 = \lim_{t' \to t^+} \min_{u_{t:t'}} \{ \frac{1}{t'-t} \int_{\tau=t}^{t'} \delta^{\tau-t} l(x,u) d\tau + \frac{\delta^{t'-t} V(x_{t'}) - V(x_t)}{t'-t} \}$$

- HJB with discount

$$-\ln(\delta) V(x_t) = \min_{u_t} \{ l(x_t, u_t) + \nabla_x V(x_t) f(x_t, u_t) \}$$

Principle of Optimality

	Finite Time	Infinite Time
Continuous Time	$-\dot{V}(t,x_t) = \min_{u_t}\{l_t(x_t,u_t) + \nabla_x V(t,x_t) f(x_t,u_t)\}$	$-\ln(\delta)V(x) = \min_{u}\{l(x,u) + \nabla_x V(x) f(x,u)\}$
Discrete Time	$V(k,x_k) = \min_{u_k}\{l_k(x_k,u_k) + V(k+1,x_{k+1})\}$	$V(x) = \min_{u}\{l(x,u) + \delta \cdot V(f(x,u))\}$

Infinite Time Optimal Control

- Stationarity
- Solving stationary Bellman equation
- Linear quadratic regulator

Bellman Operator

$$V(x) = \min_u \{l(x,u) + \delta \cdot V(f(x,u))\}$$

- Define the Bellman operator $\mathcal{T}$ as

$$\mathcal{T}(V) = \min_u \{l(x,u) + \delta \cdot V(f(x,u))\}$$

- The optimal value function is a fixed point of the Bellman operator.

- Solution strategy: value iteration

Value Iteration

- Starting from an initial estimation V_0
- Iteratively apply the Bellman operator

$$V_{i+1} = \mathcal{T}(V_i)$$

- Until convergence
- Question: is it guaranteed to converge?

Contracting Operator

- Definition: a function operator $\mathcal{F}$ is a α-contraction w.r.t. some norm $\|\cdot\|$ if

$$\forall X_1, X_2 : \|\mathcal{F}X_1 - \mathcal{F}X_2\| \leq \alpha\|X_1 - X_2\|$$

- Theorem.
  - For α-contracting operator $\mathcal{F}$ with $\alpha < 1$, there is a unique fixed point X^* which satisfies $\mathcal{F}X^* = X^*$. Moreover, all sequences $X, \mathcal{F}X, \mathcal{F}^2X, \ldots$ converge to this unique fixed point X^*.

Bellman Operator is Contracting

Proof for finite state space and L_∞ norm $\|V\|_\infty = \max_x |V(x)|$

$$\|\mathcal{T}(V_1) - \mathcal{T}(V_2)\|_\infty = \|\min_u\{l(x,u) + \delta \cdot V_1(f(x,u))\} - \min_u\{l(x,u) + \delta \cdot V_2(f(x,u))\}\|_\infty$$

$$\leq \|\max_u | \{l(x,u) + \delta \cdot V_1(f(x,u))\} - \{l(x,u) + \delta \cdot V_2(f(x,u))\} | \|_\infty$$

$$= \|\delta \max_u | V_1(f(x,u)) - V_2(f(x,u)) | \|_\infty$$

$$\leq \delta \|V_1 - V_2\|_\infty$$

Value Iteration

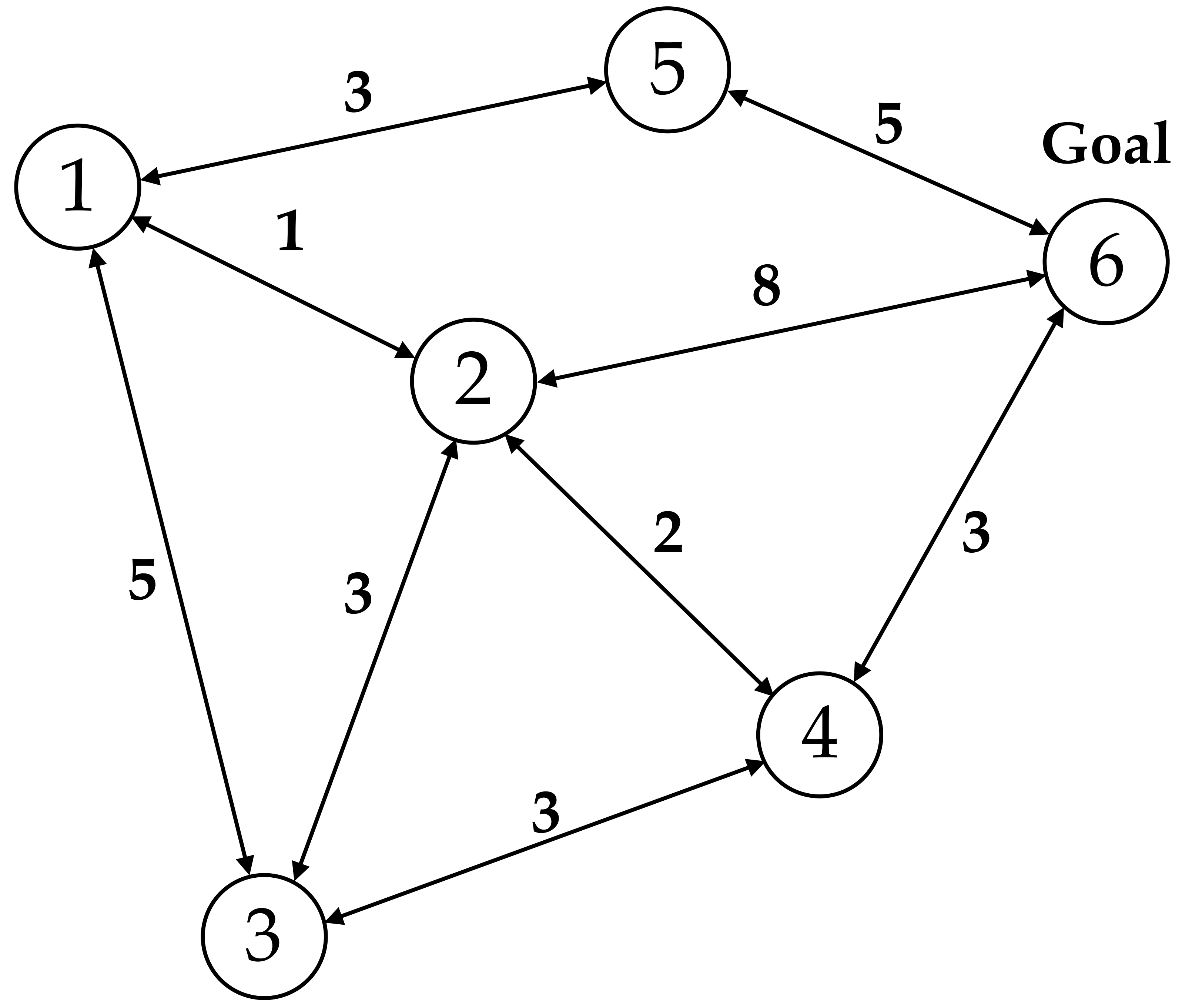
- Since the Bellman operator is contracting, value iteration will generate a converging sequence to the fixed point.
- Then the problem is how fast it can converge.
- The underlying policy may converge way before the value function converges.
- Later we will take about **policy iteration**.

Example: Shortest Path Problem

$$V(x) = \min_u \{l(x,u) + \delta \cdot V(f(x,u))\}$$

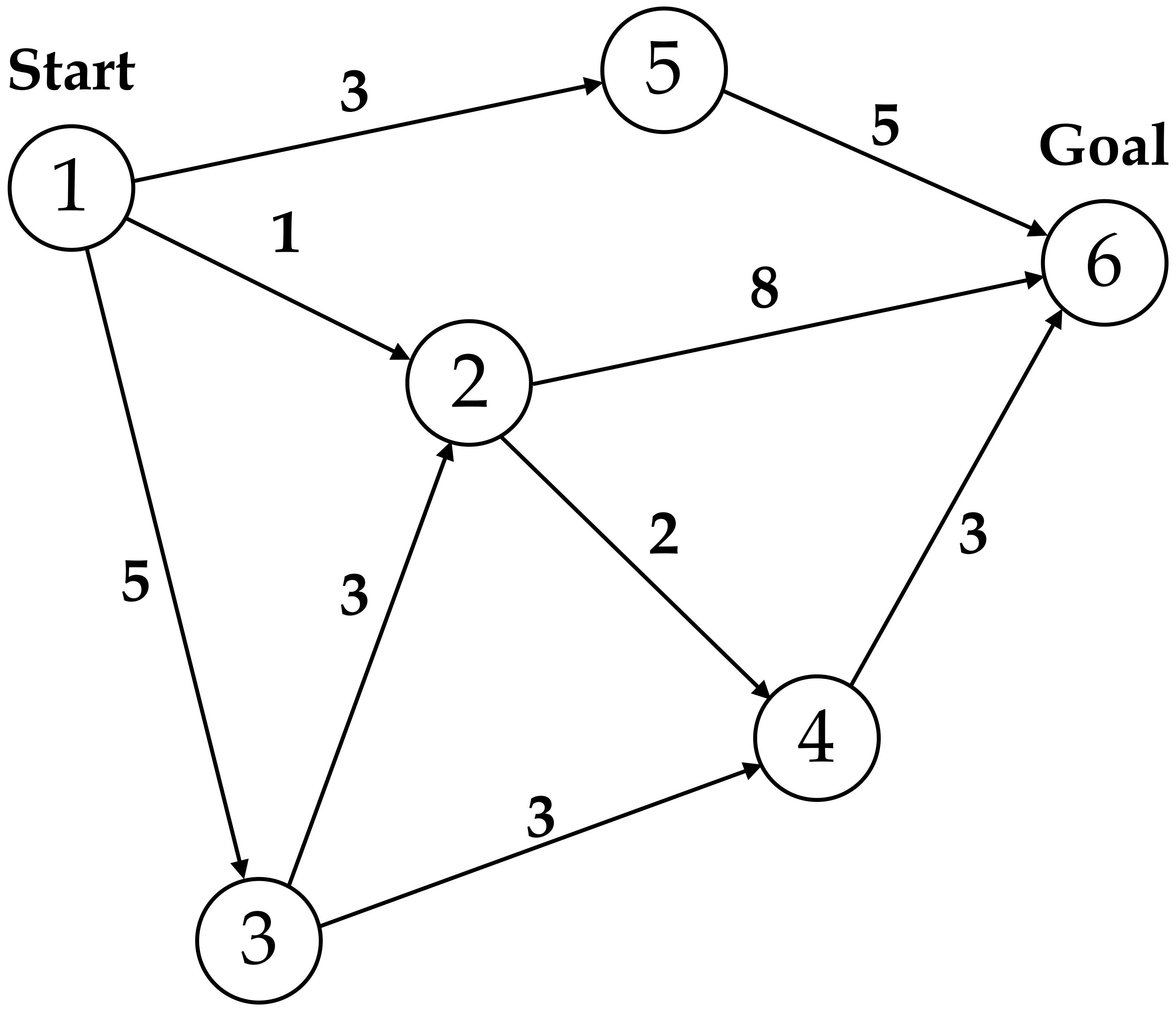
	Iteration 0	Iteration 1	Iteration 2	Iteration 3
V(1)	100	100	8 (1-5)	6 (1-2)
V(2)	100	8 (2-6)	5 (2-4)	5 (2-4)
V(3)	100	100	6 (3-4)	6 (3-4)
V(4)	100	3 (4-6)	3 (4-6)	3 (4-6)
V(5)	100	5 (5-6)	5 (5-6)	5 (5-6)
V(6)	0	0	0	0

No discount



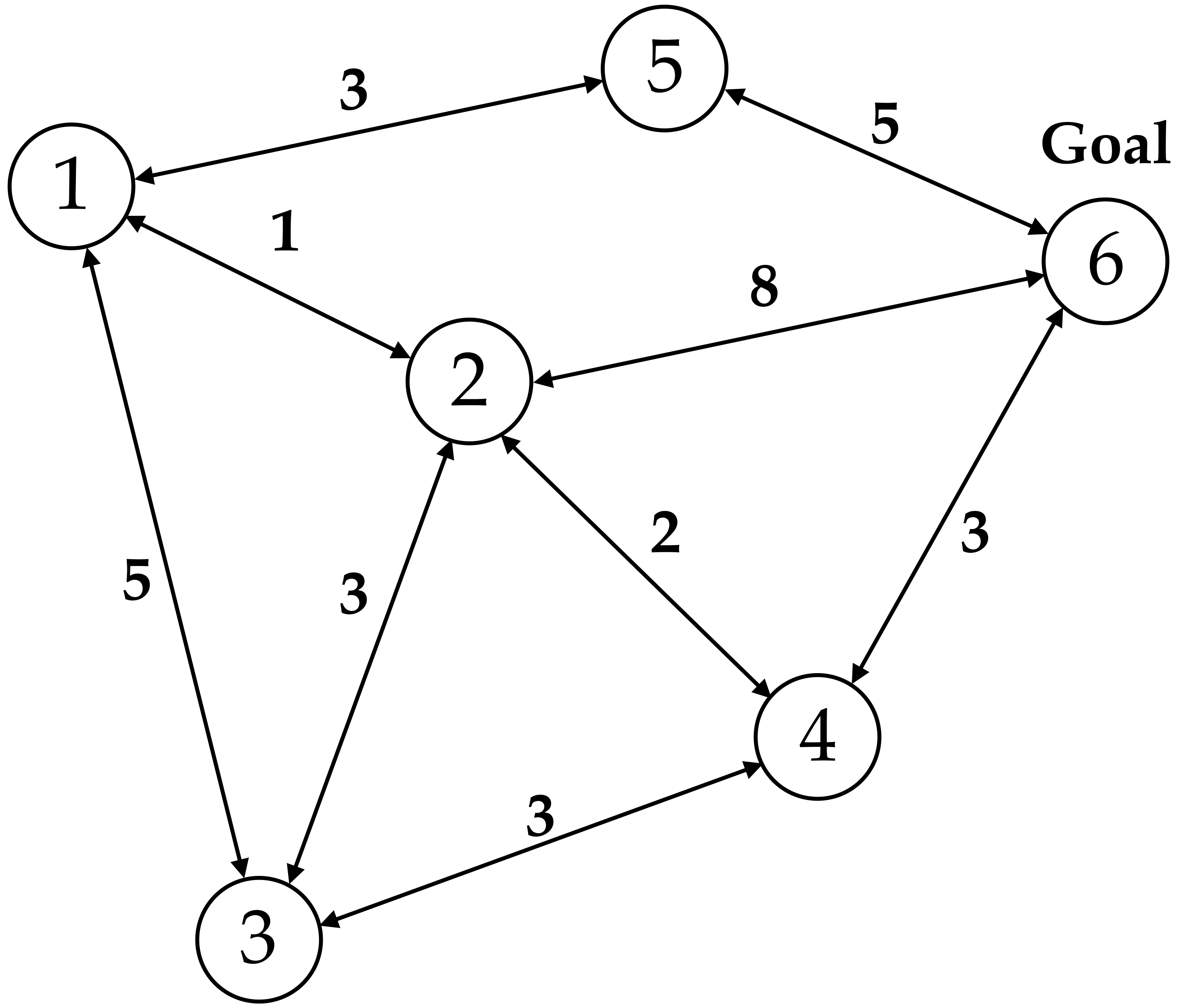
* Undirected graph, contains cycles

Example: Shortest Path Problem



* Directed graph, no cycle

Value iteration is different from backward dynamic programming that we discussed in Lecture 3!



* Undirected graph, contains cycles

Infinite Time Optimal Control

- Stationarity
- Solving stationary Bellman equation
- Linear quadratic regulator

Linear Quadratic Regulator (Infinite Continuous Time)

$$J = \frac{1}{2}\int_{t=0}^{\infty}[x^T Q x + u^T R u]dt$$

$$\dot{x} = Ax + Bu$$

Q	symmetric positive semi-definite
R	symmetric positive definite
A, B	controllable

Linear Quadratic Regulator (Infinite Continuous Time)

$$J = \frac{1}{2}\int_{t=0}^{\infty}[x^T Q x + u^T R u]dt$$

$$\dot{x} = Ax + Bu$$

Hamiltonian

$$0 = \min_u \{l(x,u) + \nabla_x V(x) f(x,u)\}$$

$$H(x,u,\lambda) = \frac{1}{2}x^T Q x + \frac{1}{2}u^T R u + \lambda^T (Ax + Bu)$$

Maximum principle

$$-\dot{\lambda} = \frac{\partial}{\partial x}H(x,u,\lambda) = Qx + A^T\lambda$$

$$0 = \frac{\partial}{\partial u}H(x,u,\lambda) = Ru + B^T\lambda \quad \Rightarrow \quad u = -R^{-1}B^T\lambda$$

Optimal Trajectory and Algebraic Riccati Equation

- According to the maximum principle, the optimal trajectory satisfies the following ODE

$$\dot{x} = Ax - BR^{-1}B^T\lambda$$
$$\dot{\lambda} = -Qx - A^T\lambda$$

- Assume that $\lambda = Px$

$$\begin{aligned} 0 &= \dot{\lambda} + Qx + A^T\lambda \\ &= P\dot{x} + Qx + A^T\lambda \\ &= PAx - PBR^{-1}B^TPx + Qx + A^TPx \end{aligned}$$

$$\boxed{PA + A^TP - PBR^{-1}B^TP + Q = 0}$$

Solving ARE

- The infinite horizon ARE

$$PA + A^T P - PBR^{-1}B^T P + Q = 0$$

is the steady state solution of the following ODE

$$\dot{P}_t + P_t A + A^T P_t - P_t BR^{-1}B^T P_t + Q = 0$$

- We can solve the infinite horizon ARE by integrating along the ODE.

- This is another form of value iteration, since the value function is

$$V(x) = \frac{1}{2}x^T Px$$

Convergence of ARE

- It remains unclear whether the "value iteration" w.r.t ARE will converge.
- We have shown that the Bellman operator is a contraction in a finite state space with a discount factor $\delta < 1$.
- Now we have infinite state space and no discount factor.

Convergence of ARE

- Proposition: If (A, B) is controllable or stablizable, then the solution to the following ODE

$$\dot{P}_t + P_t A + A^T P_t - P_t B R^{-1} B^T P_t + Q = 0$$

is guaranteed to converge to a bounded P_∞ starting from a positive definite P_0.

- Intuition: If (A, B) is not stablizable, then there are unstable states that may cause the cost function to go to infinity as t goes to infinity. It implies that the value function is unbounded, hence $V(x_0) = \frac{1}{2} x_0^T P_\infty x_0 \to \infty$ for any x_0.

Closed-Loop Property

- Proposition: The resulting closed-loop system with the optimal control $u = -R^{-1}B^T\lambda$ is asymptotically stable if (A, C) is observable or detectable for $Q = C^TC$

- Intuition: If (A, C) is not detectable, then there are unstable states that the cost function fails to penalize. In this way, the closed-loop system can be unstable but still minimizing the cost function.

Linear Quadratic Regulator (Infinite Discrete Time)

$$J = \frac{1}{2}\sum_{k=0}^{\infty}[x_k^T Q x_k + u_k^T R u_k]$$

$$x_{k+1} = Ax_k + Bu_k$$

Q	symmetric positive semi-definite
R	symmetric positive definite
A, B	controllable

Linear Quadratic Regulator (Infinite Discrete Time)

$$J = \frac{1}{2}\sum_{k=0}^{\infty}[x_k^T Q x_k + u_k^T R u_k]$$

$$x_{k+1} = Ax_k + Bu_k$$

Hamiltonian

$$V(x) = \min_u \{l(x,u) + V(f(x,u))\}$$

$$H(x_k, u_k, \lambda_{k+1}) = \frac{1}{2}x_k^T Q x_k + \frac{1}{2}u_k^T R u_k + \lambda_{k+1}^T(Ax_k + Bu_k)$$

Maximum principle

$$\lambda_k = \frac{\partial}{\partial x}H(x_k, u_k, \lambda_{k+1}) = Qx_k + A^T\lambda_{k+1}$$

$$0 = \frac{\partial}{\partial u}H(x_k, u_k, \lambda_{k+1}) = Ru_k + B^T\lambda_{k+1}$$

Discrete Algebraic Riccati Equation

- Assume that $\lambda_k = Px_k$

$$0 = Ru_k + B^T\lambda_{k+1} \quad \Rightarrow \quad u = -[B^TPB + R]^{-1}B^TPAx$$

$$\lambda_k = Qx_k + A^T\lambda_{k+1} \quad \Rightarrow \quad \boxed{P = Q + A^TPA - A^TPB[B^TPB + R]^{-1}B^TP_NA}$$

- Similar to the continuous time case, we can iteratively solve the DARE, as another form of value iteration.

- Convergence condition: (A, B) controllable or stablizable

- Closed-loop stability condition: (A, C) is observable or detectable for $Q = C^TC$

Comparison: LQRs

	Finite Time	Infinite Time
Continuous Time	$\dot{x} = A_t x + B_t u$ $J = \frac{1}{2}\int_{t=0}^{T}[x_t^T Q_t x_t + u_t^T R_t u_t]dt + \frac{1}{2}x_T^T S_T x_T$ $u_t = -R_t^{-1}B_t^T P_t x_t$ $\dot{P}_t + P_t A_t + A_t^T P_t - P_t B_t R_t^{-1} B_t^T P_t + Q_t = 0$	$\dot{x} = Ax + Bu$ $J = \frac{1}{2}\int_{t=0}^{\infty}[x^T Q x + u^T R u]dt$ $u = -R^{-1}B^T P x$ $PA + A^T P - PBR^{-1}B^T P + Q = 0$
Discrete Time	$x_{k+1} = A_k x_k + B_k u_k$ $J = \frac{1}{2}\sum_{k=0}^{N-1}[x_k^T Q_k x_k + u_k^T R_k u_k] + \frac{1}{2}x_N^T S_N x_N$ $u_k = -[R_k + B_k^T P_{k+1} B_k]^{-1} B_k^T P_{k+1} A_k x_k$ $P_k = Q_k + A_k^T P_{k+1} A_k - A_k^T P_{k+1} B_k [R_k + B_k^T P_{k+1} B_k]^{-1} B_k^T P_{k+1} A_k$	$x_{k+1} = Ax_k + Bu_k$ $J = \frac{1}{2}\sum_{k=0}^{\infty}[x_k^T Q x_k + u_k^T R u_k]$ $u = -[B^T P B + R]^{-1} B^T P A x$ $P = Q + A^T P A - A^T P B [B^T P B + R]^{-1} B^T P A$

Summary

- We have shown that the Bellman update is a contraction if
  - the state space is finite and there is a discount factor
  - the system is controllable or stablizable

Infinite Time Optimal Control

- Stationarity
- Solving stationary Bellman equation
- Linear quadratic regulator

Summary of Optimal Control

- Problems:
  - discrete time vs. continuous time
  - finite horizon vs. infinite horizon
- Bellman principle and dynamic programming
- Maximum principle and Hamiltonian
- Linear quadratic regulator and Riccati equations

Summary of Optimal Control

- All we covered so far are **exact** solutions of optimal control problems
- However, we may not always be able to get exact solution due to
  - unknown parameters in dynamics or objectives → require identification, learning and adaptation
  - computationally intractable → require approximation