



ارائه دهنده : حمید صفرزاده      آدرس ایمیل : [ha.safarzadeh@gmail.com](mailto:ha.safarzadeh@gmail.com)

**معرفی :** این درس به موضوع بهینه سازی سیستم های پویا با تکنیک برنامه ریزی پویا ( Dynamic Programming) می پردازد. در مسائل پویای مورد بررسی امکان تصمیم گیری در نقاط گسسته ای از زمان با استفاده از متغیرهای تصمیم وجود دارد و هدف بهینه سازی یک تابع هدف در افق زمانی مساله است. همچنین مسائلی که شامل تصمیم گیری چند مرحله ای باشند نیز با استفاده از این رویکرد قابل حل هستند. روش برنامه ریزی پویا به عنوان یکی از روش های اصلی بهینه سازی کاربردهای متنوعی در حوزه های مهندسی، مدیریت و اقتصاد دارد و در این درس علاوه بر بیان تئوری های مرتبط، مثال های کاربردی استفاده از این روش نیز مطرح می گردد.

**محتوای درس :** مطالب این درس در سه بخش کلی مرتبط به هم به شرح زیر ارائه می گردد:

بخش	عنوان فارسی	عنوان انگلیسی	درصد
۱	برنامه ریزی پویای قطعی	Deterministic Dynamic Programming	۱۵٪
۲	برنامه ریزی پویای تصادفی	Stochastic Dynamic Programming	۵۰٪
۳	برنامه ریزی پویای تقریبی و یادگیری تقویتی	Approximate Dynamic Programming & Reinforcement Learning	۳۵٪

**بخش ۱ -** هدف از این بخش آشنایی با ساختار یک مساله برنامه ریزی پویا و عناصر آن اعم از مرحله، حالت، تصمیم، هزینه یا پاداش و رابطه بهینگی است. با استفاده از این عناصر می توان یک مساله برنامه ریزی پویای قطعی را فرموله و حل کرد.

**بخش ۲ -** در این بخش که عمده مطالب درس را تشکیل می دهد فرض بر این است که با انتخاب مقدار متغیر تصمیم در هر مرحله، حالت مرحله بعدی و هزینه یا پاداش به وجود آمده تصادفی و با توزیع احتمال مشخص هستند. بر این اساس، هدف بهینه سازی امید ریاضی تابع هدف و همچنین تحلیل و تعیین سیاست بهینه در مساله است. مساله بیان شده با نام فرایندهای مارکوفی تصمیم گیری نیز شناخته می شود. در این بخش، مساله برنامه ریزی پویای تصادفی در دو حالت مراحل محدود و نامحدود بررسی

می‌گردد. همچنین در حالت مراحل نامحدود، مدل‌های مختلف اعم از برنامه ریزی پویا با هزینه تنزیل یافته، برنامه ریزی پویا با هزینه میانگین و برنامه ریزی پویای منفی و مثبت بررسی می‌شود.

**بخش ۳** - در این بخش با توجه به پیچیدگی محاسباتی بالای اکثر مدل‌های برنامه ریزی پویا ( curse of dimensionality) و همچنین صریح نبودن بخشی از پارامترهای مدل در برخی مسائل ( curse of modeling)، روش‌های تقریبی حل این مدل‌ها بررسی می‌شود. با توجه به اینکه مدل برنامه ریزی پویا اساس مبحث یادگیری تقویتی است، تکنیک‌های برنامه ریزی پویای تقریبی بخش مهمی از این مبحث را تشکیل می‌دهند و موضوعات ارائه شده پوشش دهنده برخی مباحث مهم یادگیری تقویتی نیز خواهد بود.

### ارزشیابی

تمرین و پروژه (۵ نمره)

آزمون میان ترم (۶ نمره)

آزمون پایان ترم (۹ نمره)

حضور در کلاس (۱ نمره اضافی)

### منابع و مراجع درس

- [1] Ross, Sheldon M. *Introduction to Stochastic Dynamic Programming*. Academic Press; 1995.
- [2] Bertsekas, Dimitri P. *Reinforcement Learning and Optimal Control*. 1st Edition. Athena Scientific; 2019.
- [3] Bertsekas, Dimitri P. *Dynamic Programming and Optimal Control*. Vol 2. 4th edition. Athena Scientific; 2012.
- [4] Puterman, Martin L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. 1st edition. Wiley-Interscience; 2005.
- [5] Powell, Warren B. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. 1st edition. Wiley-Interscience; 2007.