

علی اکبری آلاشتی

استاد بابک فرهادی

شماره دانشجویی ۴۰۲۱۲۳۳۰۱۱۷۱۶۳

### نرخ یادگیری آلفا در یادگیری تقویتی

در یادگیری تقویتی (Reinforcement Learning)، نرخ یادگیری یا آلفا ( $\alpha$ ) یکی از پارامترهای کلیدی است که تأثیر مستقیمی بر فرآیند یادگیری عامل (Agent) دارد. این پارامتر تعیین می‌کند که عامل با چه سرعتی دانش جدید را از تجربه‌های اخیر خود در مقایسه با دانش قبلی‌اش جذب می‌کند. در این متن، نقش نرخ یادگیری آلفا، نحوه عملکرد آن، تأثیراتش بر یادگیری، و نکات مربوط به تنظیم آن بررسی می‌شود.

### تعریف نرخ یادگیری آلفا

نرخ یادگیری آلفا یک مقدار عددی بین ۰ و ۱ ( $0 \leq \alpha \leq 1$ ) است که در بهروزرسانی مقادیر تابع ارزش (Value Function) یا تابع Q (Q-Function) در الگوریتم‌های یادگیری تقویتی مانند Q-Learning و SARSA استفاده می‌شود. آلفا مشخص می‌کند که چه مقدار از خطای پیش‌بینی (تفاوت بین مقدار پیش‌بینی‌شده و مقدار واقعی) باید در بهروزرسانی مقادیر قبلی تأثیر بگذارد. به عبارت دیگر، آلفا وزن اطلاعات جدید را نسبت به اطلاعات قدیمی تعیین می‌کند.

### نقش و اهمیت آلفا

نرخ یادگیری آلفا تعادل بین حفظ دانش قبلی و پذیرش اطلاعات جدید را برقرار می‌کند:

- آلفای بالا (نزدیک به ۱):

عامل بیشتر به اطلاعات جدید (تجربه‌های اخیر) وزن می‌دهد و دانش قبلی را سریع‌تر فراموش می‌کند. این حالت برای محیط‌های پویا که تغییرات زیادی دارند مناسب است، اما ممکن است باعث ناپایداری در یادگیری شود.

- آلفای پایین (نزدیک به ۰):

عامل بیشتر به دانش قبلی خود وابسته است و اطلاعات جدید تأثیر کمتری دارند. این حالت باعث یادگیری پایدارتر اما کندتر می‌شود و ممکن است در محیط‌های پویا عملکرد ضعیفی داشته باشد.

### تأثیر آلفا بر یادگیری

۱. سرعت یادگیری:

آلفای بالا باعث می‌شود عامل سریع‌تر به اطلاعات جدید واکنش نشان دهد، اما ممکن است نوسانات زیادی در مقادیر Q ایجاد کند. در مقابل، آلفای پایین یادگیری را کندتر اما پایدارتر می‌کند.

۲. همگرایی:

برای تضمین همگرایی به سیاست بهینه در الگوریتم‌هایی مانند Q-Learning، نرخ یادگیری باید به مرور زمان کاهش یابد (مثلاً با استفاده از یک برنامه کاهش آلفا). این کار باعث می‌شود عامل ابتدا به کاوش بپردازد و سپس به سمت بهره‌برداری (Exploitation) حرکت کند.

### ۳. تعادل کاوش و بهره‌برداری:

آلفای بالا می‌تواند به کاوش کمک کند، زیرا عامل به تغییرات محیط حساس‌تر است. اما اگر بیش‌ازحد بالا باشد، ممکن است مانع بهره‌برداری از دانش آموخته‌شده شود.

### تنظیم نرخ یادگیری

انتخاب مقدار مناسب برای آلفا به عوامل مختلفی بستگی دارد، از جمله پیچیدگی محیط، میزان نویز در پاداش‌ها، و نوع الگوریتم مورد استفاده. نکات زیر در تنظیم آلفا مفید هستند:

#### - مقادیر ثابت:

در مسائل ساده، می‌توان از یک مقدار ثابت (مانند  $\alpha = 0.1$ ) استفاده کرد. این روش ساده است، اما ممکن است در محیط‌های پیچیده بهینه نباشد.

#### - کاهش تدریجی آلفا:

در بسیاری از مسائل، استفاده از نرخ یادگیری کاهشی به همگرایی بهتر کمک می‌کند. این روش باعث می‌شود عامل در مراحل اولیه بیشتر کاوش کند و در مراحل بعدی به سیاست بهینه نزدیک شود.

#### - تطبیق پذیری:

در برخی الگوریتم‌های پیشرفته (مانند روش‌های مبتنی بر گرادیان)، نرخ یادگیری می‌تواند به صورت تطبیقی و بر اساس ویژگی‌های محیط تنظیم شود.

#### - آزمایش و خطا:

در عمل، تنظیم آلفا اغلب نیازمند آزمایش‌های تجربی است تا بهترین مقدار برای یک مسئله خاص پیدا شود.

### چالش‌ها و نکات

#### - نویز در محیط:

در محیط‌هایی با پاداش‌های پرنویز، آلفای پایین‌تر می‌تواند به کاهش تأثیر نویز کمک کند.

#### - محیط‌های غیرایستا:

در محیط‌هایی که قوانین یا پاداش‌ها تغییر می‌کنند، آلفای بالاتر یا روش‌های تطبیقی مناسب‌تر هستند.

#### - تأثیر بر پایداری:

آلفای بسیار بالا می‌تواند باعث نوسانات شدید در مقادیر  $Q$  شود و یادگیری را ناپایدار کند.

### نتیجه‌گیری

نرخ یادگیری آلفا یکی از مهم‌ترین پارامترها در یادگیری تقویتی است که بر سرعت، پایداری، و کیفیت یادگیری تأثیر می‌گذارد. انتخاب مقدار مناسب برای آلفا نیازمند درک عمیق از محیط و مسئله است. استفاده از نرخ‌های کاهشی یا تطبیقی می‌تواند به بهبود عملکرد الگوریتم‌های یادگیری تقویتی کمک کند. با تنظیم دقیق آلفا، می‌توان تعادل مناسبی بین کاوش و بهره‌برداری ایجاد کرد و به سیاست بهینه نزدیک شد.