۹۹۴۱۱۴۲۸ سیدپاسخ ... حمزوی

Subject: Year: Month: Day:

پاسخ تمرین کلاس 4:

الف) شرط لازم و کافی برای تضمین fixed point پاسخ را بررسی کنید و اثبات نویسید.

ب) ثابت کنید معادله بلمن این شرط را دارا است و لذا می‌توان با روش تکراری به پاسخ همگرا رسید.

الف) ابتدا لازم است تعریف فضای متریک (Metric Space) را بدانیم.

تعریف: مجموعه X که عناصر آن را نقاط می‌نامیم گفته می‌شود که یک فضای متریک است اگر هر ۲ نقطه p و q از X با یک عدد حقیقی d(p,q) مرتبط باشد که فاصله بین p و q نامیده می‌شود:

1) $d(p,q) > 0$ if $p \neq q$; $d(p,q) = 0$;

2) $d(p,q) = d(q,p)$;

3) $d(p,q) \leq d(p,r) + d(r,q)$ for any $r \in X$.

هر تابعی با 3 ویژگی بالا یک تابع فاصله (distance function) یا یک metric نامیده می‌شود.

حال به بحث تئوری پاسخ برمی‌گردیم:

بحث تئوری که وجود دارد نگاشت انقباضی است و داریم:

اگر (X,d) یک فضای متریک باشد، نگاشت $T: X \to X$ یک نگاشت انقباضی نامیده می‌شود برای X اگر یک عدد حقیقی مثبت $\alpha < 1$ وجود داشته باشد برای هر $x, y \in X$ که

$$d(Tx, Ty) \leq \alpha d(x,y)$$

از نظر هندسی، تصویر هر ۲ نقطه تحت نگاشت انقباضی نزدیک‌تر می‌شود با نسبتی که از α بزرگتر نباشد.

حالا به جزء بحث تئوری پاسخ می‌پردازیم:

King

سیده شیدا حمزوی : ۹۹۴۱۱۴۴۸

فرض کنید که X یک فضای متریک کامل متریک است و $T: X \to X$ یک نگاشت انقباضی بر روی X است، پس T دارای یک نقطه ثابت fixed point منحصر به فرد یا unique است.

اول این نشان دهیم که همیشه می‌توانیم یک $x \in X$ بیابیم که $Tx = x$ و هم یک fixed point از T است.

از یک نقطه x_0 (دلخواه) شروع می‌کنیم و سری x_k را می‌سازیم: $k = 0, 1, 2, \ldots$

$$x_k = Tx_{k-1} \quad , \quad k = 1, 2, \ldots$$

باید نشان دهیم که x_k کوشی است، در این صورت سری x_k به نقطه‌ای تحت $x \in X$ همگرا می‌شود به صورتی که X کامل است. بعد نشان می‌دهیم که $Tx = x$.

ما داریم:

$$d(x_{k+1}, x_k) \leq \alpha d(x_k, x_{k-1}) \leq \ldots \leq \alpha^k d(x_1, x_0) = \alpha^k d(x_1, x_0)$$

$$\forall k = 1, 2, \ldots$$

برای هر عدد integer m, n که $m > n$ باشد نامساوی مثلث منجر می‌شود به:

$$d(x_m, x_n) \leq \sum_{i=0}^{m-n-1} d(x_{n+i+1}, x_{n+i})$$

با ترکیب دو نامساوی بالا هم داریم:

$$d(x_m, x_n) \leq \sum_{i=0}^{m-n-1} \alpha^{n+i} d(x_1, x_0) = \alpha^n d(x_1, x_0) \frac{1 - \alpha^{m-n}}{1 - \alpha} \leq \alpha^n \frac{d(x_1, x_0)}{1 - \alpha}$$

برای هر $\varepsilon > 0$ می‌توانیم یک عدد صحیح N بیابیم که $N \geq \frac{\log \varepsilon(1-\alpha) - \log c}{\log \alpha}$ که برای هر $m, n \geq N$ داریم $d(x_m, x_n) \leq \varepsilon$

King

۹۹۴۱۱۴۸ سید بشیر خسروی

این نشان می‌دهد که سری x_n کوشی است و نقطه‌ای وجود دارد $x \in X$ که $x_n \to x$.

با توجه به نامساوی مثلث داریم:

$$d(Tx, x) \leq d(Tx, x_n) + (x_n, x) \leq \alpha d(x, x_{n-1}) + d(x_n, x)$$

$\forall k = 1, 2, \ldots$

اگر k به بی‌نهایت میل کند نامساوی بالا به:

$$d(Tx, x) = 0$$

تبدیل می‌شود. که نشان می‌دهد $Tx = x$ است یعنی x یک fixed point از T است.

فرض کنید یک fixed point دیگر به نام x' داریم:

$$d(x, x') = d(Tx, Tx') \leq \alpha d(x, x')$$

که این موضوع به وضوح به تناقض با شرط $\alpha < 1$ منجر می‌گردد پس نگاشت انقباضی T فقط یک fixed point دارد.

تا اینجا شروط بانخ و اثبات آن بیان شد.

ب) حال به سوال اصلی می‌پردازیم:

اول باید مشخص کنیم که فضای ترتیب کامل را که q_π در آن است و یک نگاشت انقباضی نیز مشخص کنیم.

فضای ترتیب کامل: در فصل فرآیندهای تصمیم مارکوف یک MDP محدود تعریف شده که S و A و R همه محدودند. فرض کنیم q_π در $\mathbb{R}^{|S| \times |A|}$ تعریف شده است. اگر $\mathbb{R}^{|S| \times |A|}$ را به همراه اندازه $\|\cdot\|$ در نظر بگیریم، به سادگی نشان داده می‌شود $\|\cdot\|$ یک norm است. یک فضای باناخ تشکیل می‌شود.

در نتیجه $(\mathbb{R}^{|S| \times |A|}, d)$ یک فضای ترتیب کامل با تابع ترتیب ناشی از اندازه (norm) است.

$$d(x, y) = \|x - y\|$$

سیده ستاره خیبری ۹۹۴۱۱۲۴۸

Subject: Year: Month: Day:

برای اثبات انقباضی بودن، معادله بلمن بهینه را در نظر بگیریم برای تابع $q: S\times A \to \mathbb{R}^{|S|\times|A|}$

تعریف می‌کنیم:

$$Tq(s,a) = \sum_{r,s'} p(r,s'|s,a)\left(r+\gamma \max_{a'} q(s',a')\right)$$

$$\forall s,s' \in S,\quad a,a' \in A$$

حال باید نشان دهیم که T یک نگاشت انقباضی است.

یک تعریف داریم که T برای یک FMDP در معادله بالا، یک نگاشت انقباضی است.

اثبات آن به صورت زیر است:

متری ترکیبی را $d(q_1,q_2)$ را در فضای $\mathbb{R}^{|S|\times|A|}$ در نظر می‌گیریم که در آن برابر است با $\|q_1-q_2\|_\infty$ برای هر $p,q \in \mathbb{R}^{|S|\times|A|}$

$$d(q_1,q_2) = \|q_1-q_2\|_\infty$$

لذا داریم:

$$\|Tq_1 - Tq_2\|_\infty = \max_{s,a} |Tq_1(s,a) - Tq_2(s,a)|$$

$$= \gamma \max_{s,a} \sum_{r,s'} p(s',r|s,a) \left|\max_{a'} q_1(s',a') - \max_{a'} q_2(s',a')\right|$$

$$\le \gamma \max_{s,a} \sum_{s'} p(s'|s,a) \max_{a'} |q_1(s',a') - q_2(s',a')|$$

$$\le \gamma \max_{s,a} \max_{s'} \max_{a'} |q_1(s',a') - q_2(s',a')|$$

$$= \gamma \max_{s',a'} |q_1(s',a') - q_2(s',a')|$$

$$= \gamma \|q_1 - q_2\|_\infty$$

پس دیدیم که عملگر بلمن بهینه انقباضی است و همچنین برای q_* می‌توان برای یک FMDP

نشان داد که بلمن اپتیمال یک و تنها یک نقطه ثابت دارد.

King