Study Report

Bae Sang Jun

UNIST

bsjuntiger@unist.ac.kr

January 8, 2025

Presentation Overview

Banach fixed-point theorem

Banach fixed-point theorem

In mathematics, the Banach fixed-point theorem is an important tool in the theory of metric spaces; it guarantees the existence and uniqueness of fixed points of certain self-maps of metric spaces, and provides a constructive method to find those fixed points.

Definition

Let (X, d) be a metric space. Then a map $T : X \to X$ is called a contraction mapping on X if there exists $q \in [0, 1)$ such that

$$d(T(x), T(y)) \leq q \cdot d(x, y)$$

for all $x, y \in X$.

Banach fixed-point theorem

Theorem

Banach Fixed-Point Theorem

Let (X, d) be a non-empty complete metric space with a contraction mapping $T : X \to X$. Then T admits a unique fixed-point x^* in X (i.e., $T(x^*) = x^*$). Furthermore, x^* can be found as follows: start with an arbitrary element $x_0 \in X$ and define a sequence $(x_n)_{n\in\mathbb{N}}$ by $x_n = T(x_{n-1})$ for $n \geq 1$. Then $\lim_{n\to\infty} x_n = x^*$.

Banach fixed-point theorem

A complete metric space is a metric space where every Cauchy sequence has a limit.

- In other words, a metric space (X, d) is complete if for every Cauchy sequence (x_n) in X, there exists $x \in X$ such that $\lim_{n \to \infty} d(x_n, x) = 0$.
- The most prominent example is the set of real numbers $\mathbb{R}$.

Banach fixed-point theorem

Proof: Existence of Fixed Point

Let's define the sequence $\{x_n\}$ as follows:

$$x_0 \in X, \quad x_{n+1} = f(x_n)$$

where x_0 is an arbitrary point in the domain of f. We will prove the existence of a fixed point by showing that $\{x_n\}$ converges to a fixed point of f. For any n, we have:

$$\begin{aligned} d(x_{n+1}, x_n) &= d(f(x_n), f(x_{n-1})) \\ &\leq k \cdot d(x_n, x_{n-1}) = k \cdot d(f(x_{n-1}), f(x_{n-2})) \\ &\leq \cdots \\ &\leq k^n \cdot d(x_1, x_0) \end{aligned}$$

Hence,

$$d(x_{n+1}, x_n) \leq kn \cdot d(x_1, x_0)$$

Therefore, $d(x_{n+1} - x_n) \leq kn \cdot d(x_1, x_0)$.

Proof: Existence of Fixed Point (Continued)

Now, let's show that $\{x_n\}$ is a Cauchy sequence. Let m and n be arbitrary natural numbers such that $m > n$. Then the following holds:

$$\begin{aligned}
d(x_m, x_n) &\leq (d(x_m, x_{m-1}) + \cdots + d(x_{n+1} - x_n)) \\
&\leq (k^{m-1} d(x_1, x_0) + \cdots + k^n d(x_1 - x_0)) \\
&= (k^{m-1} + \cdots + k^n) d(x_1 - x_0) \\
&= k^n (k^{m-n-1} + \cdots + k^0) d(x_1 - x_0) \\
&\leq k^n \left(\sum_{i=0}^{\infty} k^i \right) d(x_1 - x_0) \\
&= \frac{k^n}{1-k} d(x_1 - x_0)
\end{aligned}$$

- The first line follows from the triangle inequality.
- The second line applies the inequalities obtained for each $d(x_{i+1}, x_i)$.

As n approaches infinity, since $0 \leq k < 1$, the value of $\frac{k^n}{1-k} d(x_1 - x_0)$ converges to 0. Therefore, $d(x_m, x_n)$ also converges to 0. Hence, $\{x_n\}$ is a

Banach fixed-point theorem

Existence of Fixed Point (Continued)

Since the metric space X is complete, the Cauchy sequence $\{x_n\}$ converges to some point x. Now, let's show that x is a fixed point of f. Since f is a contraction mapping, it is Lipschitz continuous and therefore continuous. By the properties of continuous functions, we have:

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right) = f(x)$$

Also,

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n)$$

From these two equations, we can see that $x = f(x)$. Therefore, the limit point x of $\{x_n\}$ is a fixed point of f. Hence, f has at least one fixed point.

Banach fixed-point theorem

Uniqueness of Fixed Point

Now, let's prove that if a fixed point exists, it must be unique.
Let x and y be two fixed points of f. Then the following holds:

$$d(x,y) = d(f(x), f(y)) \leq k \cdot d(x,y)$$

Since $0 \leq k < 1$, the inequality above implies that $d(x,y) = 0$. By the definition of metric spaces, $x = y$. Therefore, there cannot be two distinct fixed points. Thus, the fixed point is unique.

Mutual Information

Mutual Information

Definition

Mutual Information
Consider two random variables X and Y with a joint probability mass function $p(x, y)$ and marginal probability mass functions $p(x)$ and $p(y)$. The mutual information $I(X; Y)$ is the relative entropy between the joint distribution and the product distribution $p(x)p(y)$:

$$\begin{aligned} I(X;Y) &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \\ &= D(p(x,y) \| p(x)p(y)) \\ &= E_{p(x,y)} \log \frac{p(X,Y)}{p(X)p(Y)}. \end{aligned}$$

- A measure of the amount of information that one random variable contains about another random variable.

Mutual Information

Definition

Mutual Information
Consider two random variables X and Y with a joint probability mass function $p(x,y)$ and marginal probability mass functions $p(x)$ and $p(y)$. The mutual information $I(X;Y)$ is the relative entropy between the joint distribution and the product distribution $p(x)p(y)$:

$$\begin{aligned} I(X;Y) &= \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \\ &= D(p(x,y) \| p(x)p(y)) \\ &= E_{p(x,y)} \log \frac{p(X,Y)}{p(X)p(Y)}. \end{aligned}$$

- A measure of the amount of information that one random variable contains about another random variable.

Mutual Information

Relation between Entropy and Mutual information

$$
\begin{aligned}
I(X;Y) &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \\
&= \sum_{x,y} p(x,y)\log\frac{p(x\mid y)}{p(x)} \\
&= -\sum_{x,y} p(x,y)\log p(x) + \sum_{x,y} p(x,y)\log p(x\mid y) \\
&= -\sum_{x} p(x)\log p(x) - \left(-\sum_{x,y} p(x,y)\log p(x\mid y)\right) \\
&= H(X) - H(X\mid Y).
\end{aligned}
$$

Mutual Information

$$I(X;Y) = H(X) - H(X \mid Y)I(X;X) \quad = H(X) - H(X \mid X) = H(X)$$

Thus, the mutual information of a random variable with itself is the entropy of the random variable. This is the reason that entropy is sometimes referred to as self-information.

$$
\begin{aligned}
I(X;Y) &= H(X) - H(X \mid Y) \\
I(X;Y) &= H(Y) - H(Y \mid X) \\
I(X;Y) &= H(X) + H(Y) - H(X,Y) \\
I(X;Y) &= I(Y;X) \\
I(X;X) &= H(X).
\end{aligned}
$$

Mutual Information

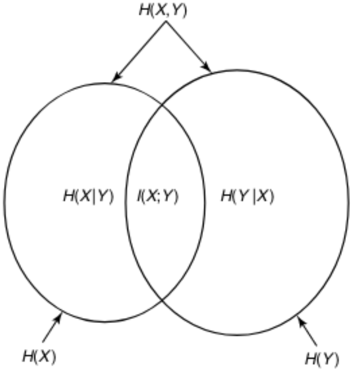

FIGURE 2.2. Relationship between entropy and mutual information.

The mutual information I(X; Y) corresponds to the intersection of the information in X with the information in Y