日本語もおｋ？

Yasunari HIKIMA

Fujitsu Limited

September 1, 2021

Uniform convergence

Definition

A sequence of functions $f_n \colon \mathbb{R} \to \mathbb{R}$ *converges uniformly* to a function $f \colon \mathbb{R} \to \mathbb{R}$ if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\sup_{x \in \mathbb{R}} |f_n(\boldsymbol{x}) - f(\boldsymbol{x})| < \epsilon.$$

Uniform convergence

Definition

A sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ *converges uniformly* to a function $f : \mathbb{R} \to \mathbb{R}$ if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\sup_{x \in \mathbb{R}} |f_n(\boldsymbol{x}) - f(\boldsymbol{x})| < \epsilon.$$

Pointwise and uniform continuity

- Uniform convergence implies pointwise convergence
- Pointwise convergence does not imply uniform convergence

Uniform convergence

Definition

A sequence of functions $f_n\colon \mathbb{R} \to \mathbb{R}$ *converges uniformly* to a function $f\colon \mathbb{R} \to \mathbb{R}$ if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\sup_{x \in \mathbb{R}} |f_n(\boldsymbol{x}) - f(\boldsymbol{x})| < \epsilon.$$

Pointwise and uniform continuity

- Uniform convergence implies pointwise convergence
- Pointwise convergence does not imply uniform convergence

Theorem

Let $f_n\colon \mathbb{R} \to \mathbb{R}$ be continuous and converge uniformly to $f\colon \mathbb{R} \to \mathbb{R}$. Then f is continuous.

Example 2 (`highlightcap`, `cbracket`)

When we consider a Gaussian prior $p(\mathbf{w}|\alpha) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \alpha^{-1}\mathbf{I})$, maximization of the corresponding posterior $p(\mathbf{w}|\mathbf{t})$ with respect to $\mathbf{w}$ is equivalent to the minimization of

$$\frac{\beta}{2}\sum_{n=1}^{N}\left\{t_n - \mathbf{w}^\top \boldsymbol{\phi}(\mathbf{x}_n)\right\}^2 + \frac{\alpha}{2}\mathbf{w}^\top\mathbf{w} \tag{3.55'}$$

the minimization corresponds to (3.27) with $\lambda = \alpha/\beta$.

Example 2 (`highlightcap`, `cbracket`)

When we consider a Gaussian prior $p(\mathbf{w}|\alpha) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \alpha^{-1}\mathbf{I})$, maximization of the corresponding posterior $p(\mathbf{w}|\mathbf{t})$ with respect to $\mathbf{w}$ is equivalent to the minimization of

$$\underbrace{\boxed{\frac{\beta}{2}\sum_{n=1}^{N}\left\{t_n - \mathbf{w}^\top \boldsymbol{\phi}(\mathbf{x}_n)\right\}^2}}_{\text{an error function}} + \frac{\alpha}{2}\mathbf{w}^\top\mathbf{w} \qquad (3.55')$$

the minimization corresponds to (3.27) with $\lambda = \alpha/\beta$.

Example 2 (`highlightcap`, `cbracket`)

When we consider a Gaussian prior $p(\mathbf{w}|\alpha) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \alpha^{-1}\mathbf{I})$, maximization of the corresponding posterior $p(\mathbf{w}|\mathbf{t})$ with respect to $\mathbf{w}$ is equivalent to the minimization of

$$\underbrace{\frac{\beta}{2}\sum_{n=1}^{N}\left\{t_n - \mathbf{w}^\top \boldsymbol{\phi}(\mathbf{x}_n)\right\}^2}_{\text{an error function}} + \underbrace{\frac{\alpha}{2}\mathbf{w}^\top\mathbf{w}}_{\text{a quadratic regularization}} \tag{3.55'}$$

the minimization corresponds to (3.27) with $\lambda = \alpha/\beta$.