Variational Inference

Lecture 2

GOALS // ○ Define Inverse Problem (IP)

○ Formulate IP via Bayes Theorem

○ Define MAP estimator via optimization

○ Define Bayes Theorem via optimization

○ variational inference.

INVERSE PROBLEM // $G : \mathbb{R}^d \to \mathbb{R}^k$, known function

Find $u \in \mathbb{R}^d$ from $y \in \mathbb{R}^k$ where

$$\boxed{y = G(u) + \eta} \qquad \text{(IP)}$$

Assumption 1 / $\eta \in \mathbb{R}^k$ is not known, but it is known to be drawn from density function ν.

○ classical approach: find point estimator u^*

○ Bayesian approach: find $\mathbb{P}(u|y)$.

BAYES THEOREM //

Assumption 2 / $u \perp\!\!\!\perp \eta$, $u \sim \rho$, $\eta \sim \nu$.

Notation / $u|y \sim \pi^y$.

Bayes Theorem / Assume that

$$Z = Z(y) := \mathbb{E}^{u \sim \rho}\, \nu(y - G(u)) > 0.$$

Then

$$\pi^y(u) = \frac{1}{Z}\, \nu(y - G(u))\, \rho(u)$$

Remark / $y|u \sim \ell(y|u) = \nu(y - G(u))$.

MAP ESTIMATION //.

Definitions / Let $L(u) = -\log \ell(y|u)$

& $R(u) = -\log \rho(u)$.

Define $J(u) = L(u) + R(u)$

Definition / A MAP estimator is point

$u_{MAP} \in \mathbb{R}^d$: $u_{MAP} \in \underset{u \in \mathbb{R}^d}{\operatorname{argmin}}\ J(u)$

Theorem / $u_{MAP} \in \underset{u \in \mathbb{R}^d}{\operatorname{argmax}}\ \pi^y(u)$

Remark: u_{MAP}: Most likely point in $\mathbb{P}(u|y)$.

VARIATIONAL BAYES // $P = P(\mathbb{R}^d)$ here.

Lemma / Let $D: P \times P \to \mathbb{R}$ be a divergence.

Then $\pi^y = \underset{q \in P}{\operatorname{argmin}} \, D(q \| \pi^y)$.

Proof / $D(\cdot \| \cdot) \geq 0$ & $D(\rho_1 \| \rho_2) = 0$ iff $\rho_1 = \rho_2$.

Example / $D_{KL}(q \| p) = \int \log\left(\frac{q(u)}{p(u)}\right) q(u) \, du$

$$= \mathbb{E}^{u \sim q} \log\left(\frac{q(u)}{p(u)}\right)$$

Definition / $J: P \to \mathbb{R}$ defined by

$$J(q) = \mathbb{E}^{u \sim q} L(u) + D_{KL}(q \| \rho).$$

Theorem / $\pi^y = \underset{q \in P}{\operatorname{argmin}} \, J(q)$

Proof / $D_{KL}(q \| \pi^y) = D_{KL}\left(q \| \frac{1}{Z} \ell(y|\cdot)\rho(\cdot)\right)$

$$= \mathbb{E}^{u \sim q} \log\left(Z q(u) / \ell(y|u)\rho(u)\right)$$

$$= \mathbb{E}^{u \sim q} \left(-\log \ell(y|u)\right) + \mathbb{E}^{u \sim q} \log\left(\frac{q(u)}{\rho(u)}\right) + \log Z$$

$$= \mathbb{E}^{u \sim q} L(u) + D_{KL}(q \| \rho) + \log Z.$$

$= J(u) + \log Z$. Apply Lemma.

VARIATIONAL INFERENCE //

Basic Idea / choose $Q \subset P$ & find

$$q^* \in \underset{q \in Q}{\arg\min}\ J(q).$$

Example / $Q = \{ q \in P : q = N(m, C)$ for some $m \in \mathbb{R}^d, C \in \mathbb{R}^{d \times d}_{\succ 0} \}$

or $Q = \{ q \in P : q = N(m, LL^T)$ for some $m \in \mathbb{R}^d, L \in \mathbb{R}^{d \times d}$ lower triangular$\}$

DATA ASSIMILATION (Smoothing) //

(a) $v^+_{j+1} = \Psi(v^+_j) + \xi^+_j, \ j \in \mathbb{Z}^+ ; \ v^+_0 \sim N(m_0, C_0)$

(b) $y^+_{j+1} = H v^+_{j+1} + \eta^+_{j+1}, \ j \in \mathbb{N}$

$\xi^+_j \sim N(0, \Sigma), \ \eta^+_{j+1} \sim N(0, \Gamma)$ i.i.d $\xi, \eta, v^+_0 \perp\!\!\!\perp$.

$$\underset{\in \mathbb{R}^{d(J+1)}}{V^+ = (v^+_0, \dots, v^+_J)} ; \quad \underset{\in \mathbb{R}^{kJ}}{Y^+ = (y^+_1, \dots, y^+_J)}$$

5

ⓐ Defines $\rho(v^+)$

ⓑ Defines $\ell(y^+ | v^+)$.

Example / MAP = w4DVar v^+_{MAP}

obtain 4DVar (strong constraint) if $\Sigma \to 0$.

Uncertainty Quantification:

$$Q = \{ q \in P(\mathbb{R}^{d(J+1)}) : q = N(v^+_{MAP}, LL^T) \quad L \text{ lower triangular} \}$$