Bilevel optimization for regression of HH onto FHN

$$ \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\norm#1{\left\lVert #1 \right\rVert} \xdef\dist{\rho} \xdef\and{\&}\xdef\AND{\quad \and \quad}\xdef\brackets#1{\left\{ #1 \right\}} \xdef\parc#1#2{\frac {\partial #1}{\partial #2}} \xdef\mtr#1{\begin{pmatrix}#1\end{pmatrix}} \xdef\bm#1{\boldsymbol{#1}} \xdef\mcal#1{\mathcal{#1}} \xdef\vv#1{\mathbf{#1}}\xdef\vvp#1{\pmb{#1}} \xdef\brackets#1{\left(#1\right)}\xdef\d{\mathrm{d}} \xdef\absval#1{\left|#1\right|} \xdef\ve{\varepsilon} \xdef\l{\lambda} \xdef\th{\vartheta} \xdef\a{\alpha} \xdef\vf{\varphi} \xdef\Tagged#1{(\text{#1})} \xdef\tagged*#1{\text{#1}} \xdef\tagEqHere#1#2{\href{#2\#eq-#1}{(\text{#1})}} \xdef\tagDeHere#1#2{\href{#2\#de-#1}{\text{#1}}} \xdef\tagEq#1{\href{\#eq-#1}{(\text{#1})}} \xdef\tagDe#1{\href{\#de-#1}{\text{#1}}} \xdef\T#1{\htmlId{eq-#1}{#1}} \xdef\D#1{\htmlId{de-#1}{\vv{#1}}} \xdef\conv#1{\mathrm{conv}\, #1} \xdef\cone#1{\mathrm{cone}\, #1} \xdef\aff#1{\mathrm{aff}\, #1} \xdef\lin#1{\mathrm{Lin}\, #1} \xdef\span#1{\mathrm{span}\, #1} \xdef\O{\mathcal O} \xdef\ri#1{\mathrm{ri}\, #1} \xdef\rd#1{\mathrm{r}\partial\, #1} \xdef\interior#1{\mathrm{int}\, #1} \xdef\proj{\Pi} \xdef\epi#1{\mathrm{epi}\, #1} \xdef\grad#1{\mathrm{grad}\, #1} \xdef\gradT#1{\mathrm{grad}^T #1} \xdef\gradx#1{\mathrm{grad}_x #1} \xdef\hess#1{\nabla^2\, #1} \xdef\hessx#1{\nabla^2_x #1} \xdef\jacobx#1{D_x #1} \xdef\jacob#1{D #1} \xdef\subdif#1{\partial #1} \xdef\co#1{\mathrm{co}\, #1} \xdef\iter#1{^{[#1]}} \xdef\str{^*} \xdef\spv{\mcal V} \xdef\civ{\mcal U} \xdef\other#1{\hat{#1}} $$

Vzpomeňme si, že SINDy metoda spočívala v optimzalizaci $$ \min _{\vv \Xi} \brackets{\frac 1 2 \norm{\vv{\dot{X}} - \vv \Theta(\vv X) \vv \Xi}_2^2 + \mu R(\vv \Xi)}, $$ což můžeme v případě LASSO regrese jakožto optimalizační metody napsat jako $$ \min _{\vv \Xi} \brackets{\frac 1 2 \norm{\vv{\dot{X}} - \vv \Theta(\vv X)\vv \Xi}_2^2 + \mu \norm{\vv \Xi}_1}. $$

Předpokládejme, že máme "pevnou" trajektorii $\vv H$ HH modelu a derivace $\vv{\dot{H}}$ a obdobně pro FHN model trajektorie $\vv F$ a příslušné derivace $\vv{\dot F}$.

Potom nalezení lineární transformace $\vv \Lambda$ modelu HH na model FHN můžeme formulovat jako úlohu $$ \min _{\vv \Lambda} \underbrace{\frac 1 2 \norm{\vv F - \vv H \vv \Lambda}_2^2}_{\mcal H(\vv \Lambda)}. \tag{\T{LTR}} $$

Jak jsme si řekli, tak uvažujeme, že trajektorie $\vv H$ je "pevná", tedy že nemůžeme měnit parametry HH modelu. Naopak o modelu FHN předpokládeme, že jeho parametry měnit můžeme. Tedy bychom chtěli nalézt pro model FHN $$ \frac{\d \vv u}{\d t} = \vv f(\vv u; \vv p), $$ kde funkce $\vv f$ zadává FHN model, $\vv p$ je vektor parametrů a $\vv u$ je stav FHN systému, takové parametry, že řeší úlohu $$ \min_{\vv p} \underbrace{\frac 1 2 \sum_{i = 1}^N \norm{\vv u(t_i; \vv p) - H_{i, \cdot} \vv \Lambda}_2^2}_{\mcal F_{\vv \Lambda}(\vv p)} \tag{\T{PRM.1}}, $$ kde $H_{i, \cdot}$ je i-té pozorování HH modelu (v čase t_i) a $\vv u(t_i; \vv p)$ je pozorování FHN modelu v čase t_i za předpokladu parametrů $\vv p$. Označme $$ \vv u(T; \vv p) = \mtr{ \vv u(t_1; \vv p) \ \vdots \ \vv u(t_N; \vv p) }, \quad T = \set{t_1, \dots, t_N}. $$ Potom můžeme tyto 2 části dát dohromady a formulovat úlohu $$ \min_{\vv \Lambda} \overbrace{\frac 1 2 \norm{\vv u(T; \vv{\hat p}) - \vv H \vv \Lambda}_2^2}^{\mcal H_{\vv{\hat p}}(\vv \Lambda)} \ \text{za podmínky } \vv{\hat p} = \argmin_{\vv p} \frac 1 2 \sum_{i = 1}^N \norm{\vv u(t_i; \vv p) - H_{i, \cdot} \vv \Lambda}_2^2 \tag{\T{F2H.1}} $$ nebo také zkrácené $$ \min_{\vv \Lambda} \mcal H_{\vv{\hat p}} (\vv \Lambda) \ \text{za podmínky } \vv{\hat p} = \argmin_{\vv p} \mcal F_{\vv \Lambda}(\vv p) $$

> Pravděpodobně nemusíme řešit případ $\vv \Lambda = \vv 0, \vv{\hat{p}} = \vv 0$, neboť i pro $\vv{\hat p} = \vv 0$ nepovoluje tvar FHN modelu konstantní nulové řešení, které by bylo best fitem pro $\vv \Lambda = \vv 0$.

Ačkoliv by tento přístup byl jistě užitečný, dostáváme me do problému s nalezením podmínky stacionarity pro optimalizaci $$ \vv{\hat p} = \argmin_{\vv p} \mcal F_{\vv \Lambda}(\vv p) $$ Pokud bychom ji chtěli najít, museli bychom spočítat $\frac {\partial \mcal F_{\vv \Lambda}}{\partial \vv p}$, avšak bez této podmínky nejsme schopni optimalizovat celou úlohu.

> Toto platí v příadě, že bychom úlohu optimalizovali metodou vyžadující gradient účelové funkce. Např. metoda "Nelder-Mead" se bez něj obejde

Proto raději použijme analogii SINDy metody, která nám umožní tento problém obejít. Proto místo úlohy $\tagEq{PRM.1}$ řešme $$ \min_{\vv p} \underbrace{\frac 1 2 \sum_{i = 1}^N \norm{\vv f(\vv H_{i, \cdot} \vv \Lambda; \vv p) - \vv{\dot{H}_{i, \cdot}} \vv \Lambda}_2^2}_{\mcal F'_{\vv \Lambda}(\vv p)}, \tag{\T{PRM.2}} $$

Revision #11
Created 28 April 2023 16:06:02 by Sceptri
Updated 6 May 2023 16:37:32 by Sceptri