

Bilevel optimization for regression of HH onto FHN

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

Vzpomeňme si, že SINDy metoda spočívala v optimalizaci $\min_{\|\mathbf{v}\|} \left\| \frac{1}{2} \|\dot{\mathbf{X}} - \mathbf{v} \Theta(\mathbf{v} \mathbf{X})\|_2^2 + \mu R(\mathbf{v}) \right\|$, což můžeme v případě LASSO regrese jakožto optimalizační metody napsat jako $\min_{\|\mathbf{v}\|} \left\| \frac{1}{2} \|\dot{\mathbf{X}} - \mathbf{v} \Theta(\mathbf{v} \mathbf{X})\|_2^2 + \mu \|\mathbf{v}\|_1 \right\|$.

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

Potom nalezení lineární transformace $\mathbf{\Lambda}$ modelu HH na model FHN můžeme formulovat jako úlohu $\min_{\|\mathbf{\Lambda}\|} \underbrace{\left\| \mathbf{F} - \mathbf{H} \mathbf{\Lambda} \right\|_2^2}_{\text{LTR}}$.

Jak jsme si řekli, tak uvažujeme, že trajektorie \mathbf{H} je "pevná", tedy že nemůžeme měnit parametry HH modelu. Naopak o modelu FHN předpokládáme, že jeho parametry měnit můžeme. Tedy bychom chtěli nalézt pro model FHN $\frac{d}{dt} \mathbf{u} = \mathbf{f}(\mathbf{u}; \mathbf{p})$, kde funkce \mathbf{f} zadává FHN model, \mathbf{p} je vektor parametrů a \mathbf{u} je stav FHN systému,

takové parametry, že řeší úlohu $\min_{\mathbf{p}} \underbrace{\frac{1}{2} \sum_{i=1}^N \|\mathbf{u}(t_i; \mathbf{p}) - \mathbf{H}_i\|_{\Lambda}^2}_{\text{mcal } F_{\Lambda}(\mathbf{p})} \tag{T\{PRM.1\}}$, kde \mathbf{H}_i je i -té pozorování HH modelu (v čase t_i) a $\mathbf{u}(t_i; \mathbf{p})$ je pozorování FHN modelu v čase t_i za předpokladu parametrů \mathbf{p} . Označme $\mathbf{u}(T; \mathbf{p}) = [\mathbf{u}(t_1; \mathbf{p}) \dots \mathbf{u}(t_N; \mathbf{p})]$, $T = \{t_1, \dots, t_N\}$. Potom můžeme tyto 2 části dát dohromady a formulovat úlohu $\min_{\Lambda} \overbrace{\frac{1}{2} \sum_{i=1}^N \|\mathbf{u}(T; \hat{\mathbf{p}}) - \mathbf{H}\|_{\Lambda}^2}^{\text{mcal } H_{\hat{\mathbf{p}}}(\Lambda)} \tag{T\{F2H.1\}}$ nebo také zkrácené $\min_{\Lambda} \text{mcal } H_{\hat{\mathbf{p}}}(\Lambda) \tag{T\{PRM.2\}}$ za podmínky $\hat{\mathbf{p}} = \argmin_{\mathbf{p}} \frac{1}{2} \sum_{i=1}^N \|\mathbf{u}(t_i; \mathbf{p}) - \mathbf{H}_i\|_{\Lambda}^2 \tag{T\{F2H.1\}}$ nebo také zkrácené $\min_{\mathbf{p}} \text{mcal } F_{\Lambda}(\mathbf{p})$

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

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

Toto platí v případě, 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 $\tag{T\{PRM.1\}}$ řešme $\min_{\mathbf{p}} \underbrace{\frac{1}{2} \sum_{i=1}^N \|\mathbf{f}(\mathbf{H}_i, \Lambda; \mathbf{p}) - \dot{\mathbf{H}}_i\|_{\Lambda}^2}_{\text{mcal } F'_{\Lambda}(\mathbf{p})} \tag{T\{PRM.2\}}$