Illesztett paraméterek hibája I

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2022. február 7.

Az illesztett paraméterek hibája

Az $y^{(i)}$ mérési pontokra illesztettünk egy $h(\mathbf{x}; \mathbf{a})$ modellt. Mennyire tekinthetőek az $\mathbf{a}$ paraméterértékek pontosnak?

Két különböző dolgot vizsgálhatunk:

- hibaterjedés: a mérési hibából következően mekkora az illesztett paraméterek bizonytalansága
- konfidencia intervallumok: mennyire lehetünk biztosak abban, hogy a mérés alapján a "valódi" paramétereket sikerült megkapni az illesztésből?

Példa: hibaterjedés egyenes illesztés esetén

$a \cdot x + b$ egyenesillesztés esetén a $y^{(i)}$ mért értékek σ_i hibája hogyan befolyásolja a kapott a és b paramétereket?

A hibaterjedés törvénye szerint egy $f(y^{(i)})$ függvény értékének hibája:

$$\sigma_f^2 = \sum_i \sigma_i^2 \left(\frac{\partial f}{\partial y^{(i)}} \right)^2$$

$\Rightarrow$ lásd Statisztikus hiba, hibaterjedés fóliák

Példa: hibaterjedés egyenes illesztés esetén

$a \cdot x + b$ egyenesillesztés esetén a $y^{(i)}$ mért értékek σ_i hibája hogyan befolyásolja a kapott a és b paramétereket?

A hibaterjedés törvénye szerint egy $f(y^{(i)})$ függvény értékének hibája:

$$\sigma_f^2 = \sum_i \sigma_i^2 \left(\frac{\partial f}{\partial y^{(i)}} \right)^2$$

$\Rightarrow$ lásd Statisztikus hiba, hibaterjedés fóliák

Emlékeztető (lásd Egyenesillesztés fóliák): $\{(x^{(i)}, y^{(i)})\}$ adathalmaz, σ_i az $y^{(i)}$ hibája. Jelölések:

$$S = \sum_i \frac{1}{\sigma_i^2} \qquad S_x = \sum_i \frac{x^{(i)}}{\sigma_i^2} \qquad S_y = \sum_i \frac{y^{(i)}}{\sigma_i^2}$$

$$S_{xx} = \sum_i \frac{(x^{(i)})^2}{\sigma_i^2} \qquad S_{xy} = \sum_i \frac{x^{(i)} y^{(i)}}{\sigma_i^2} \qquad \Delta = S \cdot S_{xx} - S_x^2$$

Hibaterjedés egyenes illesztés esetén

Ezeket felhasználva kifejezéseket kaptunk a-ra és b-re:

$$a = \frac{S_{xx} \cdot S_y - S_x \cdot S_{xy}}{\Delta} \qquad b = \frac{S \cdot S_{xy} - S_x \cdot S_y}{\Delta}$$

Ezt felhasználva:

$$\frac{\partial a}{\partial y^{(i)}} = \frac{S_{xx} - S_x\, x^{(i)}}{\sigma_i^2 \Delta} \qquad \frac{\partial b}{\partial y^{(i)}} = \frac{S\, x^{(i)} - S_x}{\sigma_i^2 \Delta}$$

a, b hibája és kovarianciája

$$\sigma_a^2 = \frac{S_{xx}}{\Delta} \qquad \sigma_b^2 = \frac{S}{\Delta} \qquad \mathrm{Cov}(a, b) = -S_x/\Delta$$

Asszimptotikus hiba

- a hibaterjedés végigszámolása bonyolult esetben nem lehetséges
- ha a modell analitikus alakja ismert, egy egyszerű módszer adható a paraméterek bizonytalanságának becslésére

Tekintsük $\chi^2(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ viselkedését a minimum körül:

- $\chi^2(\mathbf{a})$ az $\mathbf{a}_0$ minimum körül Taylor-sorba fejthető

$$\chi^2(\mathbf{a}) = \chi_0^2 + \left.\frac{\partial \chi^2}{\partial a_k}\right|_{\mathbf{a}_0} (\mathbf{a} - \mathbf{a}_0) + (\mathbf{a} - \mathbf{a}_0)^\mathsf{T} \, \frac{1}{2} \left.\frac{\partial^2 \chi^2}{\partial a_k \partial a_l}\right|_{\mathbf{a}_0} (\mathbf{a} - \mathbf{a}_0) + \ldots$$

- a minimumhelyen az első derivált definíció szerint 0
- a második derivált pozitív definit, az *irányonkénti* nagysága jellemzi, hogy “mennyire stabil” a minimum
- bízunk benne, hogy a magasabb rendű tagok kicsik

A Hesse-mátrix

Egy többváltozós függvény "görbületét" jellemzi a parciális második deriváltak segítségével.

$\chi^2(\mathbf{a})$-re felírva:

$$2 \cdot \alpha_{kl} = \frac{\partial^2 \chi^2(\mathbf{a})}{\partial a_k \partial a_l}$$

Állítás:

- az α_{kl} mátrix **inverze** jellemzi az illesztett paraméterek standard hibáját
- az átlós elemekben σ_k^2 jelenik meg
- a nem diagonális elemekben a k. és l. paraméterek kovarianciája
- normál eloszlású mérési hibák és lineáris illesztés esetén ez egzaktul belátható
- egyéb esetben csak (jó) közelítés

A Hesse-mátrix általános lineáris függvényillesztés esetén

Az egyszerűség kedvéért tegyük fel, hogy a mérés hibája állandó: $\sigma_i = \sigma$

$\chi^2(\mathbf{a})$ és a szükséges deriváltak kifejezése a $\mathbf{X}$ tervmátrix segítségével:

$$\begin{aligned}
\chi^2(\mathbf{a}) &= \sum_i \left[\sum_j (X_{ij} a_j) - y^{(i)}/\sigma \right]^2 \\
\frac{\partial \chi^2(\mathbf{a})}{\partial a_k} &= 2 \cdot \sum_i \left[\sum_j X_{ik} X_{ij} a_j - X_{ik} y^{(i)}/\sigma \right] \\
\frac{\partial^2 \chi^2(\mathbf{a})}{\partial a_k \partial a_l} &= 2 \cdot \sum_i \left[X_{ik} X_{il} \right]
\end{aligned}$$

A Hesse-mátrix általános lineáris függvényillesztés esetén

Az egyszerűség kedvéért tegyük fel, hogy a mérés hibája állandó: $\sigma_i = \sigma$

$\chi^2(\mathbf{a})$ és a szükséges deriváltak kifejezése a $\mathbf{X}$ tervmátrix segítségével:

$$
\begin{aligned}
\chi^2(\mathbf{a}) &= \sum_i \left[\sum_j (X_{ij} a_j) - y^{(i)}/\sigma \right]^2 \\
\frac{\partial \chi^2(\mathbf{a})}{\partial a_k} &= 2 \cdot \sum_i \left[\sum_j X_{ik} X_{ij} a_j - X_{ik} y^{(i)}/\sigma \right] \\
\frac{\partial^2 \chi^2(\mathbf{a})}{\partial a_k \partial a_l} &= 2 \cdot \sum_i \left[X_{ik} X_{il} \right]
\end{aligned}
$$

A Hesse és a kovariancia mátrix

$$
\boldsymbol{\alpha} = \mathbf{X}^\mathsf{T} \mathbf{X} \qquad \mathbf{C} = \boldsymbol{\alpha}^{-1}
$$

Példa: parabola illesztése 5 pontra

A mérési adatok

$$x^{(i)} = \{ \quad -2 \quad , \quad -1 \quad , \quad 0 \quad , \quad 1 \quad , \quad 2 \quad \}$$
$$y^{(i)} = \{ \quad 5.1, \quad 1.9, \quad 1.1, \quad 2.1, \quad 4.9 \}$$

A mérési hiba az egyszerűség kedvéért $\sigma_i = \sigma = 0.1$

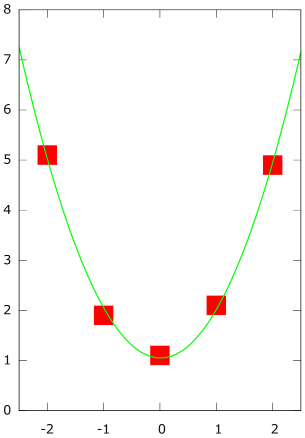
A modell: $f(x) = a_0 + a_1 x + a_2 x^2$, azaz $f_1(x) = 1$, $f_2(x) = x$ és $f_3(x) = x^2$.

$$\mathbf{X} = 10 \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \qquad \mathbf{X}^{\mathrm{T}}\mathbf{X} = 10^2 \begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix}$$

A függvényillesztés eredménye

Az $\mathbf{X}^\mathsf{T}\mathbf{X}\mathbf{a} = \mathbf{X}^\mathsf{T}\mathbf{y}$ egyenletet **a**-ra megoldva:

$$\mathbf{a}_{fit} = \begin{bmatrix} 1.049 \\ -0.020 \\ 0.986 \end{bmatrix}$$



A konkrét példában

Az illesztett modell:

$$f(x) = 1.049 - 0.020x + 0.986x^2$$

A kovarianciamátrix:

$$(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1} = 10^{-2}\begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix}^{-1} = 10^{-2}\begin{bmatrix} 0.49 & 0 & -0.14 \\ 0 & 0.10 & 0 \\ -0.14 & 0 & 0.07 \end{bmatrix}$$

Vagyis az egyes paraméterek szórása és kovarianciája:

$$\begin{aligned} \sigma(a_0) &= 0.070 \\ \sigma(a_1) &= 0.032 \\ \sigma(a_2) &= 0.027 \\ \mathrm{cov}(a_0, a_2) &= -0.014 \end{aligned}$$

Feladat: számoljuk ki a Hesse mátrixot és az inverzét az egyenesillesztés esetére! Látunk-e különbséget a kapott hibában a hibaterjedés képletének felhasználásával kapott eredményhez képest?