Illesztett paraméterek hibája I

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2023. március 2.

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?

Hibaterjedés

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

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

ahol σ_i az $y^{(i)}$ hibája
$\Rightarrow$ lásd Statisztikus hiba, hibaterjedés fóliák

Hibaterjedés

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

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

ahol σ_i az $y^{(i)}$ hibája
$\Rightarrow$ lásd Statisztikus hiba, hibaterjedés fóliák

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

- erre az esetre explicit ki lehet számolni az a és b hibáját a fenti egyenlet segítségével
- lásd Egyenesillesztés fóliák

Asszimptotikus hiba

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

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 $\sigma^2(a_k)$ 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

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?

A Hesse-mátrix általános lineáris függvényillesztés esetén

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

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

A Hesse-mátrix általános lineáris függvényillesztés esetén

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

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

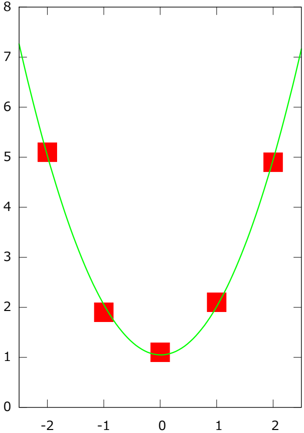
A Hesse és a kovariancia mátrix

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

Példa: parabola illesztése 5 pontra

Az illesztett modell:

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



Példa: parabola illesztése 5 pontra

A kovarianciamátrix:

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

Az egyes paraméterek szórása és kovarianciája:

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

A $\sigma(a_i)$, $i = 0, 1, 2$ jellemzi az egyes paraméterek meghatározásának hibáját