Adatmodellezés
Lineáris függvényillesztés

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2020. február 13.

Elméleti modell

Fizikai törvény:

- Egy elmélet, valamilyen idea a világról
- *Matematikai összefüggéseket* adunk a mennyiségek között
- Az összefüggések *konstansokat*, *paramétereket* tartalmaznak

Példa: feldobott kő

$$F = m\ddot{x} = -g \cdot m$$

A differenciálegyenlet megoldása:

$$x(t) = -\frac{g}{2}t^2 + v_0 t + x_0$$

Elméleti modell

Fizikai törvény:

- Egy elmélet, valamilyen idea a világról
- *Matematikai összefüggéseket* adunk a mennyiségek között
- Az összefüggések *konstansokat*, *paramétereket* tartalmaznak

Példa: feldobott kő

$$F = m\ddot{x} = -g \cdot m$$

A differenciálegyenlet megoldása:

$$x(t) = -\frac{g}{2}t^2 + v_0 t + x_0$$

Az elméleti modell bizonyos x_i *változók* és **a** *paramétervektor* függvényében *becslést ad* a mérhető fizikai mennyiségek értékeire.

Elméleti modell ellenőrzése

Kísérlet:

- bizonyos *mennyiségeket mérünk*
- sokszor más mennyiségek függvényében, pl. az idő vagy hely

Problémák:

- a mért adatok sosem pontosan a valóságot adják vissza
- az adatokat torzítás, zaj és mérési hiba terheli

Elméleti modell ellenőrzése

Kísérlet:

- bizonyos *mennyiségeket mérünk*
- sokszor más mennyiségek függvényében, pl. az idő vagy hely

Problémák:

- a mért adatok sosem pontosan a valóságot adják vissza
- az adatokat torzítás, zaj és mérési hiba terheli

Cél:

- ellenőrizni a fizikai törvény helyességét,
- megadni a konstansok és paraméterek értékét,
- továbbá megmondani, hogy a kísérlet alapján mennyire lehetünk abban bizonyosak, hogy a megalkotott törvény a valóságot írja le

Melyik a legjobban illeszkedő modell?

A modell által adott $y(x_i|\mathbf{a})$ becslések jól kell, hogy illeszkedjenek az y_i mért értékekre.

A mért y_i értékek:

- a mérésben mindig csak valamilyen $P(y_i)$ valószínűséggel fordulnak elő
- a $P(y_i)$ valószínűség eloszlása megismételt mérésekkel elvileg megbecsülhető
- a mérési hibáról magáról is felteszünk valamit, ez is a modell része
- ha $P(y_i)$ normális eloszlásnak felel meg, akkor a mérési hibának a σ szórás vehető

Melyik a legjobban illeszkedő modell?

A modell által adott $y(x_i|\mathbf{a})$ becslések jól kell, hogy illeszkedjenek az y_i mért értékekre.

A mért y_i értékek:

- a mérésben mindig csak valamilyen $P(y_i)$ valószínűséggel fordulnak elő
- a $P(y_i)$ valószínűség eloszlása megismételt mérésekkel elvileg megbecsülhető
- a mérési hibáról magáról is felteszünk valamit, ez is a modell része
- ha $P(y_i)$ normális eloszlásnak felel meg, akkor a mérési hibának a σ szórás vehető

Hogyan számszerűsíthető, hogy mennyire illeszkedik jól a modell a mérésekhez?

A legkisebb négyzetek módszere

Definiáljuk az ún. *khi-négyzet költségfüggvényt*:

$$\chi^2 = \sum_i \frac{\left[y_i - y(x_i|\mathbf{a})\right]^2}{\sigma_i^2}$$

- ez fogja jellemezni az modell illeszkedésének jóságát
- általában a σ_i szórás (mérési hiba) minden mérési pontban különbözhet
- azok a mérési pontok, ahol σ_i nagy, kisebb súllyal szerepelnek

A legkisebb négyzetek módszere

Definiáljuk az ún. *khi-négyzet költségfüggvényt*:

$$\chi^2 = \sum_i \frac{[y_i - y(x_i|\mathbf{a})]^2}{\sigma_i^2}$$

- ez fogja jellemezni az modell illeszkedésének jóságát
- általában a σ_i szórás (mérési hiba) minden mérési pontban különbözhet
- azok a mérési pontok, ahol σ_i nagy, kisebb súllyal szerepelnek

Miért pont a χ^2 költségfüggvényt használjuk?

- ha $P(y_i)$ normális eloszlású σ_i szórással, akkor megmutatható, hogy a *maximum likelihood becslés* a χ^2 költségfüggvényhez vezet

Költségfüggvény:

$$\chi^2 = \sum_i \frac{[y_i - y(x_i|\mathbf{a})]^2}{\sigma_i^2}$$

legkisebb négyzetek módszere: a modell **a** paraméterereire úgy kaphatunk becslést, ha minimalizáljuk a χ^2-t:

$$\arg\min_a \sum_i \frac{[y_i - y(x_i|\mathbf{a})]^2}{\sigma_i^2} = ?$$

A legkisebb négyzetek módszere

Megjegyzés: eddig a modellről nem tettünk fel semmit:

- lehet matematikai formula
- lehet algoritmus **a** bemenő paraméterekkel

A legkisebb négyzetek módszere

Megjegyzés: eddig a modellről nem tettünk fel semmit:

- lehet matematikai formula
- lehet algoritmus **a** bemenő paraméterekkel

Ha a model egy matematikai formula, akkor a minimumban:

$$\frac{\partial \chi^2(\mathbf{a})}{\partial a_k} = 0$$

vagyis a χ^2 minimalizációs feltétel:

$$0 = -2 \sum_i \frac{[y_i - y(x_i|\mathbf{a})]}{\sigma_i^2} \frac{\partial y(x_i|...a_k...)}{\partial a_k}$$

minden k-ra, ahol k a paraméterek számáig futó index.

A legkisebb négyzetek módszere

Megjegyzés: eddig a modellről nem tettünk fel semmit:

- lehet matematikai formula
- lehet algoritmus **a** bemenő paraméterekkel

Ha a model egy matematikai formula, akkor a minimumban:

$$\frac{\partial \chi^2(\mathbf{a})}{\partial a_k} = 0$$

vagyis a χ^2 minimalizációs feltétel:

$$0 = -2 \sum_i \frac{[y_i - y(x_i|\mathbf{a})]}{\sigma_i^2} \frac{\partial y(x_i|...a_k...)}{\partial a_k}$$

minden k-ra, ahol k a paraméterek számáig futó index.

Általános esetben a χ^2 bonyolult kifejezés.

- ha ismert is zárt alakban, analitikusan nehéz kezelni
- a parciális deriváltakat nem feltétlenül könnyű kiszámolni
- előfordulhat, hogy több lokális minimuma van, nehéz megtalálni az abszólút minimumot (ha létezik)

A legkisebb négyzetek módszere

További megjegyzések:

- a számolás során x_i-ket végig ismertnek vettük, de általános esetben ezeknek is lehet hibájuk
- ha a modell egy algoritmus **a** bemenő paraméterekkel, akkor a minimum keresés is csak algoritmikusan kezelhető

Példa: egyenes illesztése

Ismert:

- $x_1, x_2, ..., x_i$ mérési pontok, ezeknek nincs hibájuk
- $y_1, y_2, ..., y_i$ mért értékek
- $\sigma_1, \sigma_2, ..., \sigma_i$ becsült hibák

Feladat: illesszünk a pontokra egyenest χ^2 módszerrel.

- modell: $y(x) = y(x|a, b) = a + bx$

Az optimalizálandó költségfüggvény:

$$\chi^2 = \sum_i \left(\frac{y_i - y(x_i|a, b)}{\sigma_i} \right)^2 = \sum_i \left(\frac{y_i - a - bx_i}{\sigma_i} \right)^2$$

Milyen a és b mellett lesz χ^2 minimális?

Példa: egyenes illesztése

A minimumhelyen a parciális deriváltak eltűnnek:

$$
\begin{aligned}
0 &= \frac{\partial \chi^2}{\partial a} = -2 \sum_i \frac{y_i - a - bx_i}{\sigma_i^2} \\
0 &= \frac{\partial \chi^2}{\partial b} = -2 \sum_i \frac{x_i(y_i - a - bx_i)}{\sigma_i^2}
\end{aligned}
$$

Megjegyzések:

- ez egy lineáris egyenletrendszer a-ra és b-re
- biztos, hogy minimumhelyet találunk, mert χ^2 kifejezése pozitív kvadratikus.

Példa: egyenes illesztése

Jelölések:

$$S = \sum_i \frac{1}{\sigma_i^2} \qquad S_x = \sum_i \frac{x_i}{\sigma_i^2} \qquad S_y = \sum_i \frac{y_i}{\sigma_i^2}$$

$$S_{xx} = \sum_i \frac{x_i^2}{\sigma_i^2} \qquad S_{xy} = \sum_i \frac{x_i y_i}{\sigma_i^2}$$

Példa: egyenes illesztése

Jelölések:

$$S = \sum_i \frac{1}{\sigma_i^2} \qquad S_x = \sum_i \frac{x_i}{\sigma_i^2} \qquad S_y = \sum_i \frac{y_i}{\sigma_i^2}$$

$$S_{xx} = \sum_i \frac{x_i^2}{\sigma_i^2} \qquad S_{xy} = \sum_i \frac{x_i y_i}{\sigma_i^2}$$

Az új jelölésekkel az egyenletrendszer a-ra és b-re:

$$\begin{aligned} Sa + S_x b &= S_y \\ S_x a + S_{xx} b &= S_{xy} \end{aligned}$$

Példa: egyenes illesztése

Az egyenletrendszer determinánsa:

$$\Delta = S \cdot S_{xx} - S_x^2$$

Az egyenletrendszer megoldása:

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

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

Kilógó pontok

Egy mérés során becsúszhatnak rossz mérések

- ezeket nem jellemzi a mérési hiba
- valamilyen ritka, nem várt esemény hatására

A modellillesztés során a kilógó pontoktól érdemes valamilyen módon megszabadulni.

Kilógó pontok

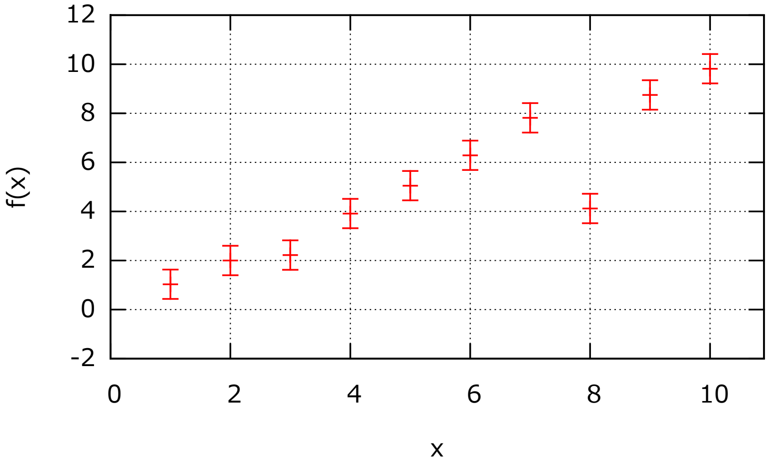


Figure: Lineáris függés, kilógó ponttal

Kilógó pontok kezelése

Léteznek kifinomult becslő módszerek, ezek pl akkor alkalmazhatóak, ha ismert, hogy mi a mérési hiba eloszlása: $\Rightarrow$ "**Robust estimation**"

Kilógó pontok kezelése

Léteznek kifinomult becslő módszerek, ezek pl akkor alkalmazhatóak, ha ismert, hogy mi a mérési hiba eloszlása: $\Rightarrow$ "**Robust estimation**"

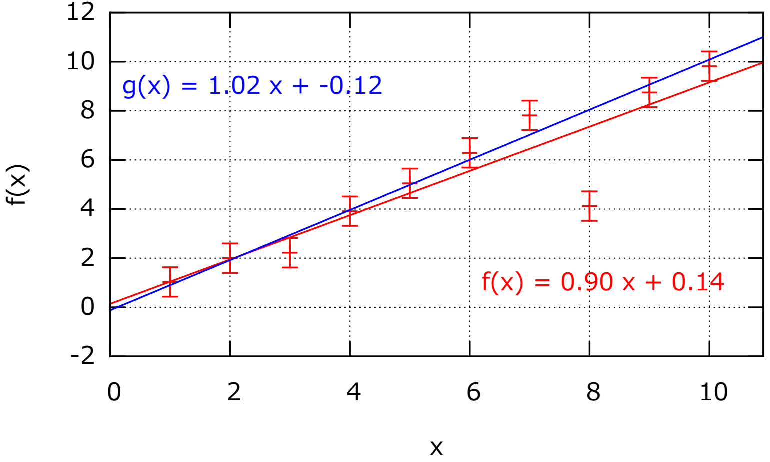
Helyette egy egyszerű iteratív módszer:

- illesszük a modellt a mérési pontokra
- számoljuk ki a mérési pontoktól vett eltérések szórását
- dobjuk ki azokat a pontokat, amik 3σ-n kívül esnek
- ismételjük meg az illesztést

A módszer kevés kilógó ponttal elbánik

- arra számítunk, hogy egy idő után nem lesz 3σ-n kívüli pont

Kilógó pontok



Általános lineáris függvényillesztés

Az $(x_i,\ y_i)$ adathalmazra szeretnénk illeszteni egy modellt. Mint az előzőekben, most is a

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - f(x_i|\mathbf{a})}{\sigma_i} \right]^2$$

összeget kell minimalizálni.

$f(x|\mathbf{a})$-t a következő alakban keressük:

$$f(x|\mathbf{a}) = \sum_{j=1}^{M} a_j f_j(x),$$

ahol $f_j(x)$ tetszőleges ún. *bázisfüggvény*, ami már nem függ az a_j-ktől.

A probléma lineáris, mert a teljes $f(x|\mathbf{a})$ a különböző $f_j(x)$-ek lineárkombinációja.

Általános lineáris függvényillesztés

A minimumot a

$$\frac{\partial \chi^2(a_j)}{\partial a_j} = 0$$

feltétel adja.

Általános lineáris függvényillesztés

A minimumot a

$$\frac{\partial \chi^2(a_j)}{\partial a_j} = 0$$

feltétel adja.

Behelyettesítve $f(x|\mathbf{a})$-t:

$$\frac{\partial \chi^2(a_j)}{\partial a_j} = 2 \cdot \sum_{i=1}^{N} \left[\frac{1}{\sigma_i^2} \cdot \left(\sum_{k=1}^{M} a_k f_k(x_i) - y_i \right) \cdot f_j(x_i) \right] = 0$$

Ez egy lineáris egyenletrendszer az a_k együtthatókra!

A tervmátrix[1]

Az átláthatóság kedvéért vezessük be a következőket:

$$X_{ij} = \frac{f_j(x_i)}{\sigma_i} \qquad\qquad b_i = \frac{y_i}{\sigma_i}$$

X_{ij} az úgynevezett *tervmátrix*:

- az M oszlopa a bázisfüggvényeknek felel meg
- az N sora a mérési pontoknak
- a mátrixelemek a j. bázisfüggvény x_i helyeken vett értékei

[1] design matrix

A tervmátrix és a b_i vektor felépítése

$$X_{ij} = \begin{pmatrix} \frac{f_1(x_1)}{\sigma_1} & \frac{f_2(x_1)}{\sigma_1} & \dots & \frac{f_M(x_1)}{\sigma_1} \\ \frac{f_1(x_2)}{\sigma_2} & \frac{f_2(x_2)}{\sigma_2} & \dots & \frac{f_M(x_2)}{\sigma_2} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \frac{f_1(x_N)}{\sigma_N} & \frac{f_2(x_N)}{\sigma_N} & \dots & \frac{f_M(x_N)}{\sigma_N} \end{pmatrix} \qquad b_i = \begin{pmatrix} \frac{y_1}{\sigma_1} \\ \frac{y_2}{\sigma_2} \\ \vdots \\ \vdots \\ \frac{y_N}{\sigma_N} \end{pmatrix}$$

A lineáris illesztés normálegyenletei

A parciális deriváltakra felírt egyenletek ezzel:

$$\sum_{i=1}^{N} \left[\left(\sum_{k=1}^{M} a_k X_{ik} - b_i \right) \cdot X_{ij} \right] = 0,$$

ami indexes írásmóddal:

$$X_{ik} a_k X_{ij} = X_{ij} b_i$$

A lineáris illesztés normálegyenletei

A parciális deriváltakra felírt egyenletek ezzel:

$$\sum_{i=1}^{N}\left[\left(\sum_{k=1}^{M} a_k X_{ik} - b_i\right)\cdot X_{ij}\right] = 0,$$

ami indexes írásmóddal:

$$X_{ik} a_k X_{ij} = X_{ij} b_i$$

Mátrixos írásmódban:

$$\mathbf{X}^{\mathsf{T}}\mathbf{X}\mathbf{a} = \mathbf{X}^{\mathsf{T}}\mathbf{b},$$

ahol **X** egy $N \times M$-es mátrix és **a** tartalmazza az ismeretleneket.

A lineáris illesztés normálegyenletei

A parciális deriváltakra felírt egyenletek ezzel:

$$\sum_{i=1}^{N}\left[\left(\sum_{k=1}^{M} a_k X_{ik} - b_i\right)\cdot X_{ij}\right] = 0,$$

ami indexes írásmóddal:

$$X_{ik} a_k X_{ij} = X_{ij} b_i$$

Mátrixos írásmódban:

$$\mathbf{X}^{\mathsf{T}}\mathbf{X}\mathbf{a} = \mathbf{X}^{\mathsf{T}}\mathbf{b},$$

ahol **X** egy $N \times M$-es mátrix és **a** tartalmazza az ismeretleneket.

Ez egy $M \times M$ lineáris egyenletrendszer: a legkisebb négyzetes illesztés **normálegyenletei**

Megoldás: pl Gauss-Jordan elimináció

Többváltozós polinomillesztés

Ha az y_i mérési pontokat nem egyetlen x_i változó függvényében mérjük, hanem az x_i mérési pontok maguk is $x_i^{(k)}$ K-dimenziós vektorok

Ekkor definiáljuk a következő polinomot

$$f(\mathbf{x}|\mathbf{a}) = \sum_{k=1}^{K} \sum_{j=1}^{M} a_j^{(k)} [x^{(k)}]^{j-1},$$

- a probléma immár összesen $M \cdot K$ ismeretlent fog tartalmazni
- ebben nincsenek vegyes tagok!
- a probléma ugyanúgy oldható meg, mint az előző

További fontos kérdések

A paraméterek meghatározásával még nem ért véget a feladat:

- mekkora a meghatározott paraméterek hibája?
- egyáltalán mennyire jó a modell? Hiába kicsi a meghatározott paraméterek hibája, ha rossz a modell, amit használunk

Az illesztett paraméterek hibája

Elvégeztünk egy optimalizációs eljárást, ami a mérési adatok alapján megadott bizonyos **a** modellparamétereket. Mennyire tekinthetőek ezek 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 megilleszteni?

A paraméterek hibája egyenes illesztés esetén

Az y_i mért értékek hibája mennyire befolyásolja a kapott a és b paraméterek bizonytalanságát?

A hibaterjedés törvénye szerint egy függvény értékének hibája:

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

A paraméterek hibája egyenes illesztés esetén

Az y_i mért értékek hibája mennyire befolyásolja a kapott a és b paraméterek bizonytalanságát?

A hibaterjedés törvénye szerint egy függvény értékének hibája:

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

A fentiekben explicit kifejezéseket kaptunk a-ra és b-re. Ezt felhasználva:

$$\frac{\partial a}{\partial y_i} = \frac{S_{xx} - S_x x_i}{\sigma_i^2 \Delta} \qquad \frac{\partial b}{\partial y_i} = \frac{S x_i - S_x}{\sigma_i^2 \Delta}$$

Behelyettesítve:

$$\sigma_a^2 = \frac{S_{xx}}{\Delta} \qquad \sigma_b^2 = \frac{S}{\Delta}$$

Kiszámolható a kovariancia is: $\mathrm{Cov}(a, b) = -S_x/\Delta$

Asszimptotikus hiba

Általánosabb esetben, ha a modell analitikus alakja ismert, egy egyszerű módszer a paraméterek bizonytalanságára.

Tekintsük χ^2 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

M változó esetén parciális második deriváltakat kell nézni: *Hesse-mátrix*

A Hesse-mátrix

Egy többváltozós függvény "görbületét" jellemzi. Írjuk fel a χ^2-re:

$$2 \cdot \alpha_{kl} = \frac{\partial^2 \chi^2}{\partial a_k \partial a_l}$$

Állítás:

- az α_{kl} mátrix **inverze** jellemzi az illesztett paraméter 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, nem lineáris függvényillesztés esetén csak (jó) közelítés.

Egyenesillesztés hibája a Hesse-mátrix inverzével

A χ^2 parciális deriváltjait a és b szerint már kiszámoltuk a minimum keresésekor:

$$\frac{\partial \chi^2}{\partial a} = -2 \sum_i \frac{y_i - a - bx_i}{\sigma_i^2} \qquad \frac{\partial \chi^2}{\partial b} = -2 \sum_i \frac{x_i(y_i - a - bx_i)}{\sigma_i^2}$$

A második parciális deriváltakból alkotott Hesse-mátrix a korábbi jelölésekkel:

$$2 \cdot \alpha = \begin{pmatrix} \frac{\partial^2 \chi^2}{\partial a^2} & \frac{\partial^2 \chi^2}{\partial a \partial b} \\ \frac{\partial^2 \chi^2}{\partial b \partial a} & \frac{\partial^2 \chi^2}{\partial b^2} \end{pmatrix} = 2 \cdot \begin{pmatrix} S & S_x \\ S_x & S_{xx} \end{pmatrix}$$

Az α mátrixot invertálva kapjuk a hibákat és a kovarianciákat:

$$\alpha^{-1} = \begin{pmatrix} \sigma_a^2 & \mathrm{Cov}(a, b) \\ \mathrm{Cov}(a, b) & \sigma_b^2 \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} S_{xx} & -S_x \\ -S_x & S \end{pmatrix} \checkmark$$

A Hesse-mátrix általános lineáris függvényillesztés esetén

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

Az $\mathbf{X}$ tervmátrix segítségével: $(\mathbf{Xa} - \mathbf{y}/\sigma)_i = \sum_j X_{ij} a_j - y_i/\sigma$

A Hesse-mátrix általános lineáris függvényillesztés esetén

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

Az $\mathbf{X}$ tervmátrix segítségével: $(\mathbf{Xa} - \mathbf{y}/\sigma)_i = \sum_j X_{ij} a_j - y_i/\sigma$

Ennek segítségével a χ^2 és a szükséges deriváltak:

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

Az illesztett paraméterek hibája

A második parciális deriváltakból álló Hesse-mátrix:

$$\boldsymbol{\alpha} = \mathbf{X}^\mathsf{T}\mathbf{X}$$

Ennek az inverze adja a kovarianciamátrixot:

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

- az átlós elemek a varianciákat tartalmazzák: $\sigma_k^2 = C_{kk}$
- a többi a kovarianciákat: $\mathrm{Cov}(k, l) = C_{kl}$

Példa: parabola illesztése 5 pontra

Mivel öt megadott pont esetében a parabolaillesztés túlhatározott, a legkisebb négyzetek módszerét használjuk. Legyenek a mérési adatok a következők (valójában oszlopvektorok):

$$\begin{array}{l} \mathbf{x} = \{ \quad -2 \quad , \quad -1 \quad , \quad 0 \quad , \quad 1 \quad , \quad 2 \quad \} \\ \mathbf{y} = \{ \quad 5.1, \quad 1.9, \quad 1.1, \quad 2.1, \quad 4.9 \} \end{array}$$

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

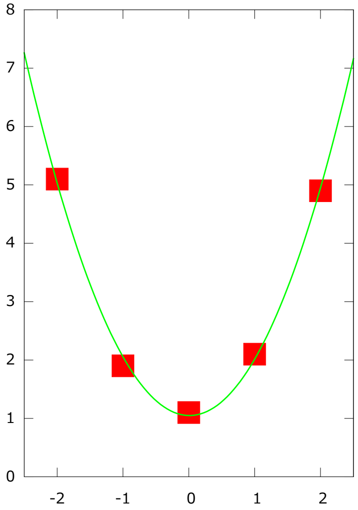
A modell három ismeretlenes: $f(x) = a_0 + a_1 x + a_2 x^2$, azaz $f_1(x) = 1$, $f_2(x) = x$ és $f_3(x) = x^2$. Ezzel:

$$\mathbf{X} = 10 \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \mathbf{X}^\mathsf{T}\mathbf{X} = 10^2 \begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix} \quad \mathbf{X}^\mathsf{T}\frac{\mathbf{y}}{\sigma} = 10^2 \begin{bmatrix} 15.1 \\ -0.2 \\ 44.0 \end{bmatrix}$$

A függvényillesztés eredménye

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

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



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:

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