# Lineáris függvényillesztés

2019. február 11.

# Elméleti modell

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

A modell sokféle lehet:

- tisztán matematikai: függvényillesztés
  ilyenkor az $x_i$ értékeket adottak, és az $a$ paramétereket kell variálni úgy, hogy a modell által adott $y(x_i|a)$ becslések jól illeszkedjenek az $y_i$ mért értékekre.

$$y(x_i|a) = f(x_i, a)$$

- szimulációs: numerikus algoritmus, stb. (ld. később)

Kérdések:

- mennyire jó egy modell?
- a mérések alapján melyik a legjobb modell?
- hogyan találjuk meg a legjobb modellt?

# Melyik a legjobban illeszkedő modell?

Ha a modell szerint a mérési eredmények egymástól *függetlenek*, és kizárólag az $x_i$ értékektől és az ismeretlen $a$ paraméterektől függenek, akkor használhatjuk a *legnagyobb valószínűség*[1] módszerét.

A mért $y_i$ értékek sosem pontosak. Konkrét mérés esetében egy adott mért érték mindig csak valamilyen $P(y_i)$ valószínűséggel fordul elő.

- a hibáról magáról is felteszünk valamit, ez is a modell része
- a $P(y_i)$ valószínűség eloszlása megismételt mérésekkel elvileg megbecsülhető
- standard hiba esetében ez normális eloszlás a mérési hibának megfelelő $\sigma$ szórással

[1] angolul: maximum likelihood

# A *likelihood*-függvény

Kérdés: Mi annak a valószínűsége, hogy a modell egy adott $a$ paraméterezése mellett a méréssorozat pont a mért $y_i$ értékeket adja?

Legyen $p(y_i|x_i, a)$ annak a valószínűségnek az eloszlása, hogy egy konkrét paraméterezés esetén az $i$. mérés éppen $y_i$ értéket ad.

Ha a méréssorozat $N$ független mérés egymásutánjából áll össze, akkor a méréssorozat megvalósulásának teljes valószínűségi eloszlása adott $a$ paraméterek esetén:

$$L(a) = \prod_i^N p(y_i|x_i, a)$$

Ez az ún. *likelihood-függvény*, aminek a maximumát keressük az $a$ paraméterek variálása mellett.

# A likelihood-függvény normális eloszlású hiba esetén

Tegyük fel, hogy $p(y_i|x_i, a)$ a Gauss-eloszlás, azaz a mért érték normális eloszlású a modell által becsült érték körül $\sigma_i$ szórással:

$$p(y_i|x_i, a) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y_i - y(x_i|a)}{\sigma_i}\right)^2\right]$$

Ezzel pedig a likelihood-függvény:

$$L(a) = \prod_i \left\{\frac{1}{\sigma_i\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y_i - y(x_i|a)}{\sigma_i}\right)^2\right]\right\}$$

Keressük a likelihood-függvény maximumát, azaz:

$$\arg\max_a L(a) = ?$$

# A likelihood-függvény logaritmusa

$L(a)$ kifejezésében a produktum hasában csak pozitív számok állnak. Vegyük az egész kifejezés logaritmusát. Mivel a logaritmus monoton, $\ln L$ maximuma ugyanott lesz, ahol $L$ maximuma.

$$\ln L(a) = \sum_i \left[ -\frac{1}{2} \left( \frac{y_i - y(x_i|a)}{\sigma_i} \right)^2 \right] + C,$$

ahol a $C$ konstans tartalmaz minden olyan tagot, ami nem függ az $a$ paraméterektől.

Definiáljuk a következő mennyiséget:

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

Ez az ún. *khi-négyzet*, amiről látszik, hogy a minimuma pontosan olyan $a$ paramétereknél van, ahol $L(a)$ maximális.

# A legkisebb négyzetek módszere

Konstans, normális eloszlású hiba esetén $\sigma$ kiemelhető a $\chi^2$ kifejezéséből, így – végső soron – a maximum likelihood módszer átmegy a *legkisebb négyzetek módszerébe*:

$$\arg\min_a \sum_i \left[y_i - y(x_i|a)\right]^2 = ?$$

A számolás során $x_i$-ket végig ismertnek vettük.

- Általános esetben ezeknek is lehet hibájuk.

Eddig a modellről nem tettünk fel semmit:

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

# A $\chi^2$ minimalizálása

Általános esetben a $\chi^2$ bonyolult kifejezés.

- ha ismert is zárt alakban, analitikusan nehéz kezelni
- ha a modell csak algoritmikusan adott, akkor a minimum is csak algoritmikusan kezelhető

Függvények minimumának keresésére majd nézünk módszereket.

Pár érdekes eset:

- a függvény kvadratikus
- ha a függvénynek lokális minimumai vannak
- ha nagyon sok lokális minimum van

# Kvadratikus kifejezés minimuma

Mivel ez egy négyzetes kifejezés, jó eséllyel lézezik minimumhelye-
Ebben a pontban az $a$ szerinti parciális deriváltak értéke 0. A konkrét esetben

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

a következőre vezet:

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

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

Kérdés persze, hogy a parciális deriváltakat mennyire könnyű kiszámolni.

# Az illesztett paraméterek hibája

Eddig arról volt szó, hogy a mért értékeknek van valamilyen hibájuk.

Most: elvégeztünk egy optimalizációs eljárást, ami a mérési adatok alapján megadott bizonyos *a* modellparamétereket. Kérdés, hogy ezek a paraméterértékek mennyire tekinthetők 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?
  szemléletesen: mennyire romlik el az illesztés, ha a legjobban illeszkedő paramétereket egy kicsit megvariáljuk

# Illesztett paraméterek hibája

Eddig azt néztük, hogy a mérési hiba hogyan propagál az illesztett paraméterekbe, ha van egy konkrét függvényünk.

- a hibaterjedés végigszámolása bonyolult esetben nem lehetséges
- főleg, ha a modell nem analitikus
- a kovarianciákat nehéz megbecsülni

Mégis hogyan lehet jellemezni az illesztett paraméterek megbízhatóságát?
Két lehetőség adódik:

- az illesztett modellparamétereket kicsit megváltoztatva mennyire rontjuk el az illesztést
- a mérést az illesztett paraméterekkel „újraszimulálva”, majd „újraillesztve” mennyire szórnak az illesztett értékek

# Az illesztés jósága: redukált $\chi^2$

A $\chi^2$ definíció szerint:

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

Ez függ a mérési pontok számától, ezért két méréssorozat illesztésének jóságát nem fogjuk tudni összehasonlítani vele.

Kérdés: miként kell normálni?

$N$ a mérési pontok száma, és $M$ az illesztett paraméterek száma

- definiáljuk a szabadsági fokok számát: $\nu = N - M$

# Redukált $\chi^2$

- ha az illesztett modell az *a* paraméterekben lineáris, akkor belátható, hogy $\chi^2$ eloszlása a minimum körül azonos az ún. $\nu$ szabadsági fokú $\chi^2$ eloszlással
- ez $\nu$ egymástól független normális eloszlású véletlen változó összegének eloszlása, aminek a várható értéke is $\nu$
- normáljunk tehát a szabadsági fokok számával!

Normáljuk $\chi^2$-et $\nu = N - M$-mel, a szabadsági fokok számával. Ekkor az „egész jó” illesztés feltétele:

$$\frac{\chi^2}{\nu} \approx 1$$

# Aszimptotikus hiba

Egyszerű módszer a paraméterek bizonytalanságának becslésére, ha a modell analitikus alakja ismert.

Tekintsük $\chi^2$ viselkedését a minimum körül:

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

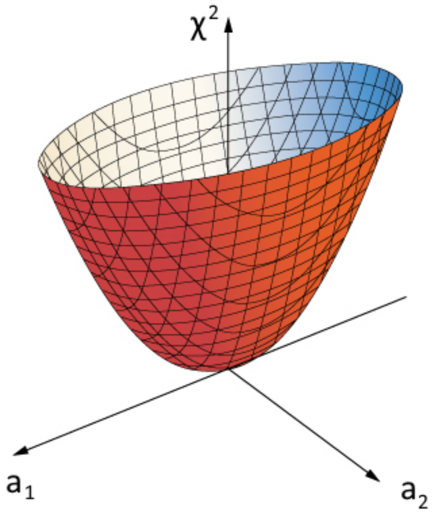
$$\chi^2(\mathbf{a}) = \chi_0^2 + \left.\frac{\partial \chi^2}{\partial a_k}\right|_{\mathbf{a}_0} (\mathbf{a} - \mathbf{a}_0) + (\mathbf{a} - \mathbf{a}_0)^\mathsf{T} \left.\frac{\partial^2 \chi^2}{\partial a_k \partial a_l}\right|_{\mathbf{a}_0} (\mathbf{a} - \mathbf{a}_0) + \dots$$

- 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

# Pozitív definit kvadratikus kifejezés



# A Hesse-mátrix inverze

Egy többváltozós függvény „görbületét" jellemzi. Írjuk fel a $\chi^2$-re:

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

Állítás:

- az $\alpha_{kl}$ mátrix *inverze* jellemzi az illesztett paraméter standard hibáját
- az átlós elemekben $\sigma_k^2$ 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.

# A paraméterek hibájának becslése más módon

A Hesse-mátrix csak analitikus modell esetében használható. Más modellek esetében próbálkozhatunk numerikus deriválással.

Egy jellegében más módszer: nézzük az illesztett paraméterek stabilitását úgy, hogy variáljuk a pontokat, amikre illesztünk

- erre a nagy ágyú módszer a Monte Carlo, de az bonyolult

Két egyszerűbb módszer:

- Jackknife módszer
- Bootstrapping

# Jackknife[2] módszer

Tekintsük a mérési pontokat, de minden lépésben hagyjunk ki egyet az illesztésből

- hagyjuk ki az $i$. pontot
- illesszük a modellt $N - 1$ pontra
- legyen az illesztett paraméterek vektora $\theta_i$

Minden egyes mérési pontra megismételve összesen $N$ különböző paramétervektort kapunk

- ezek átlaga lesz a becsült paramétervektor

$$\theta_{\mathrm{jack}} = \frac{1}{N} \sum_i \theta_i$$

- ezek szórása az illesztett paraméterek standard hibája

$$\sigma^2_{\mathrm{jack}} = \frac{N-1}{N} \sum_i (\theta_i - \theta_{\mathrm{jack}})^2$$

[2] jackknife = bicska

# Bootstrapping

Most válasszunk ki véletlenszerűen valamennyit az eredeti mérési pontokból, és illesszünk csak azokra

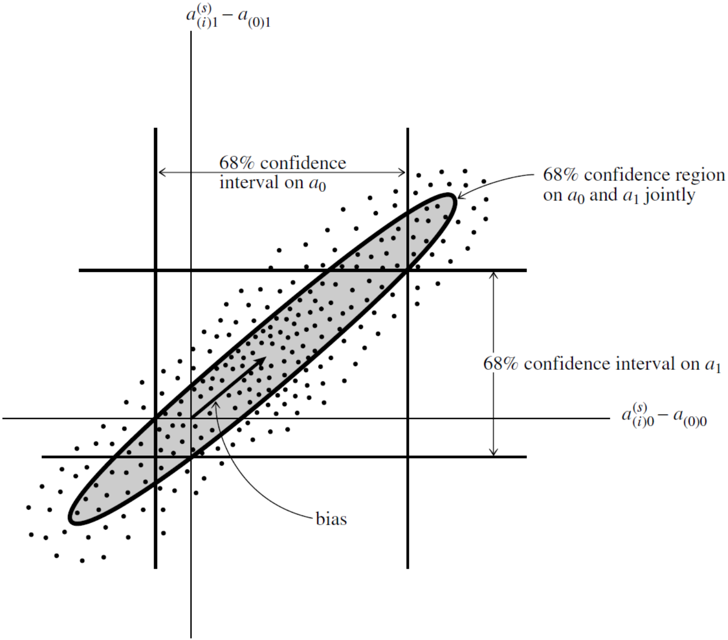
- ez elvileg $\binom{n}{k}$ módon megtehető, de nem kell ennyit végignézni

Az eljárás ugyanaz, mint az előző

- meghatározzuk a $\theta_i$ paramétervektorokat
- kiszámítjuk az átlagot és a szórást minden paraméterre

Mindkét módszer esetében számolhatjuk a paraméterek kovarianciáját is.

# Konfidenciatartományok



# Kilógó pontok

Egy mérés során becsúszhatnak rossz mérések

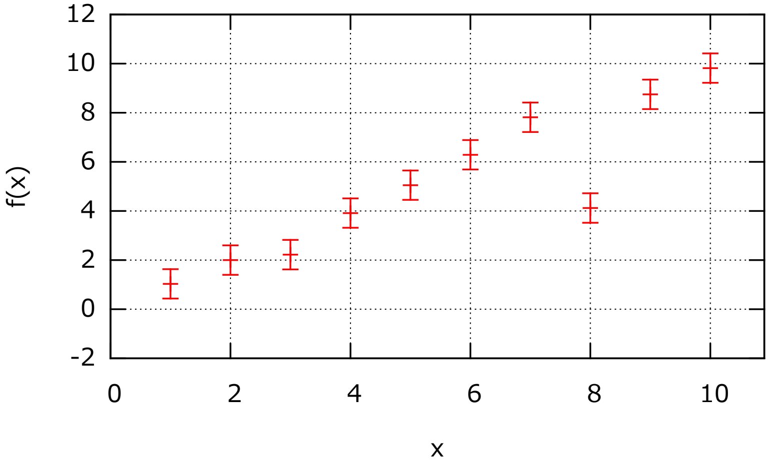
- ezeken nem jellemzi a mérési hiba
- valami miatt nem a mérési előírás szerint mérünk
- valamilyen ritka, nem várt esemény hatására

A modellillesztés során a kilógó pontoktól érdemes valamilyen módon megszabadulni.

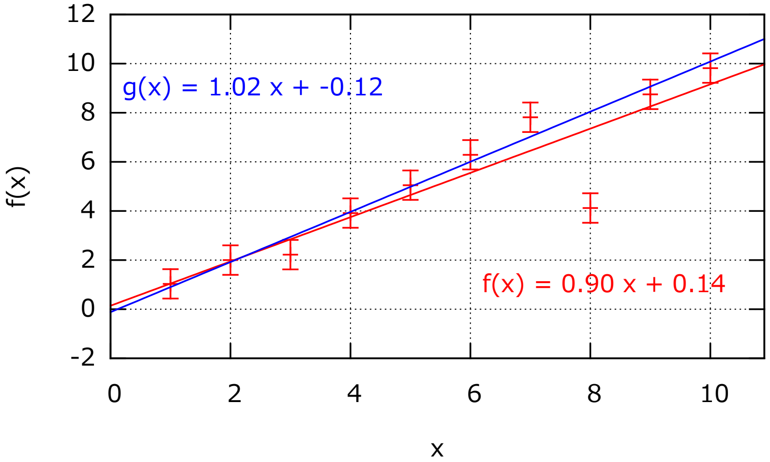
Példa:

- egy távcső CCD detektorát a fotometriai mérés során egy nagy energiájú kozmikus részecske éri el, és nagy számú extra elektront gerjeszt

# Kilógó pontok



# Kilógó pontok



# Kilógó pontok kezelése

Léteznek szofisztikált robusztus becslő módszerek, ezek bonyolultak.

- ezek a mérések hibáját nem normális eloszlásúnak, hanem hosszú farkúnak tételezik fel

Helyette nézzünk egy gyakori iteratív módszert

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

A módszer kevés kilógó ponttal elbánik

- arra számítunk, hogy egy idő után nem lesz $3\sigma$-n kívüli pont

# 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 $\chi^2$ 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 $\chi^2$ minimális?

# A minimum megkeresése

A költségfüggvény:

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

A minimumhelyen a parciális deriváltak eltűnnek:

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

Tudnivalók:

- Ez egy lineáris egyenletrendszer $a$-ra és $b$-re, csak ki kell bogarászni az együtthatókat.
- Biztos, hogy minimumhelyet találunk, mert $\chi^2$ kifejezése pozitív kvadratikus.

# Új jelölések

Jelölések:

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

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

# Lineáris egyenletrendszer $a$-ra és $b$-re

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

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

Az egyenletrendszer determinánsa:

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

Az egyenletrendszer megoldása:

$$\begin{aligned} a &= \frac{S_{xx} \cdot S_y - S_x \cdot S_{xy}}{\Delta} \\ b &= \frac{S \cdot S_{xy} - S_x \cdot S_y}{\Delta} \end{aligned}$$

# A megoldás hibájának meghatározása

Itt most azt nézzük, hogy az $y_i$ mért értékek hibája mennyire befolyásolja a kapott $a$ és $b$ paraméterek bizonytalanságát.

A hibaterjedés törvénye szerint egy függvény értékének hibája:

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

Nekünk most az $a$ és $b$ paraméterek hibáját kell tekintenünk az $y_i$ mért értékek függvényében, tehát:

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

Behelyettesítve:

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

Kiszámolható a kovariancia is: $\mathrm{cov}_{ab} = -S_x/\Delta$

# A Hesse-mátrix inverzével

A $\chi^2$ 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 $\alpha$ mátrixot invertálva kapjuk a hibákat és a kovarianciákat:

$$\alpha^{-1} = \begin{pmatrix} \sigma_a^2 & \mathrm{cov}_{ab} \\ \mathrm{cov}_{ab} & \sigma_b^2 \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} S_{xx} & -S_x \\ -S_x & S \end{pmatrix}$$

# Általános lineáris függvényillesztés

Nem csak az egyenesillesztés lineáris probléma. Keressük a modellt az általános $f(x|a_j)$ alakban.

- $x$ lehet vektor is, de most nem írunk neki indexet
- ha $x$ indexet kap, az a mérési pontokra vonatkozik majd
- $a$ az $M$ darab illesztendő paraméter
- ekkor $\chi^2 = \chi^2(a)$ is függeni fog a paraméterektől
- és persze $y_i$-ktől és $\sigma_i$-ktől is

Tekintsük a $\chi^2$ parciális deriváltjaira a minimumban teljesülő feltételeket.

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

# Elvégezve a parciális deriválást

$\chi^2$ korábbi definíciója alapján a deriváltakra adódó feltétel:

$$\frac{\partial \chi^2}{\partial a_j} = 2 \cdot \sum_{i=1}^{N} \left[ \frac{f(x_i|a) - y_i}{\sigma_i^2} \cdot \left. \frac{\partial f(x|a)}{\partial a_j} \right|_{x=x_i} \right] = 0$$

Eddig a pontig nem tettünk fel semmit $f(x|a)$ konkrét alakjáról

- nem is tudunk egyszerű megoldást adni
- a $\frac{\partial f(x|a)}{\partial a_j}$ derivált lehet nagyon bonyolult

Ötlet: korlátozzuk a problémát olyan esetekre, amikor a parciális derivált nem függ az $a_j$ paraméterektől.

# A lineáris probléma

Keressük $f(x|a)$-t a következő alakban:

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

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

- A probléma lineáris, mert a teljes $f(x|a)$ a különböző $f_j(x)$-ek lineárkombinációja.
- Az $f_j(x)$-ek konkrét alakja tetszőleges.

Így már el tudjuk végezni a parciális deriválást:

$$\frac{\partial f(x|a)}{\partial a_k} = f_k(x),$$

hiszen

$$\frac{\partial (a_j f_j(x))}{\partial a_k} = \delta_{jk} f_j(x)$$

# A megoldandó lineáris egyenletrendszer

Behelyettesítve az $f$ tetszőleges függvények lineárkombinációjából álló alakját a következő adódik:

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

A szögletes zárójelen belüli szorzat harmadik tagjában azért nincsen már $\sum$, mert a Kronecker-$\delta$ elvitte a többi tagot.

Vegyük észre, hogy ez már lineáris egyenletrendszer az $a_k$ együtthatókra.

# A tervmátrix[3]

A probléma innentől már csak egy lineáris egyenletrendszer megoldása. 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

[3] 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:

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

Kibogozva az indexet a következő mátrixegyenletet kapjuk:

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

ahol **X** egy $N \times M$-es mátrix, tehát végső soron $M$ darab $M$ ismeretlenes egyenletet kell megoldani, például Gauss-eliminációval.

# Az Hesse-mátrix

$$\begin{aligned}
(\mathbf{Xa} - \mathbf{y})_i &= \sum_j x_i^{(j-1)} a_j - y_i \\
\chi^2 &= \sum_i \left[ \sum_j \left( x_i^{(j-1)} a_j \right) - y_i \right]^2 \\
\frac{\partial \chi^2}{\partial a_k} &= 2 \cdot \sum_i \left[ \sum_j x_i^{(k-1)} x_i^{(j-1)} a_j - x_i^{(k-1)} y_i \right] \\
\frac{\partial^2 \chi^2}{\partial a_k \partial a_l} &= 2 \cdot \sum_i \left[ x_i^{(k-1)} x_i^{(l-1)} \right]
\end{aligned}$$

# Az illesztett paraméterek hibája

A második parciális deriváltakból álló Hesse-mátrix valójában egyszerűen

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

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: $\mathrm{cov}_{kl} = C_{kl}$

Az illesztés aszimptotikus hibájának általában a redukált $\chi^2$-tel szorzott varianciát vesszük.

# Többváltozós polinomillesztés

Ha az $x_i$ mérési pontok maguk is $x_i^{(k)}$ K-dimenziós vektorok, akkor a többdimenziós polinomillesztés modellje a következő:

$$f_j^{(k)}(x) = \sum_{p=1}^{M \cdot K} a_p x^{j-1},$$

- a probléma immár összesen $M \cdot K$ ismeretlent fog tartalmazni
- de ebben még nincsenek vegyes tagok
- a $p$ index a $j$ és $k$ indexekből képzett rendezett párok halmazán (Descartes-szorzat) fut
- a probléma ugyanúgy oldható meg, mint az előző
- a végén $M \cdot K \times M \cdot K$ méretű mátrixot kell invertálni

# 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):

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

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:

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

# A függvényillesztés eredménye

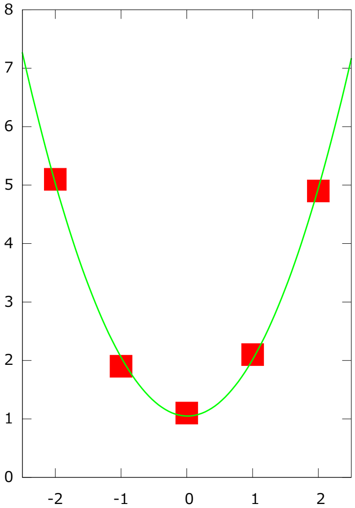
Az $\mathbf{X}^\mathrm{T}\mathbf{X}\mathbf{a} = \mathbf{X}^\mathrm{T}\mathbf{y}$ egyenletet $\mathbf{a}$-ra megoldva:

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

Az illesztés jósága:

$$\begin{aligned} \chi^2 &= (\mathbf{X}\mathbf{a} - \mathbf{y})^2 \\ &= 0.0411 \end{aligned}$$

$$\frac{\chi^2}{\mathrm{NDF}} = \frac{0.0411}{5-3} = 0.0206$$



# A konkrét példában

Az illesztett modell:

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

A kovarianciamátrix:

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

Vagyis az egyes paraméterek szórása és kovarianciája:

$$\begin{aligned} \sigma_0 &= 0.70 \\ \sigma_1 &= 0.32 \\ \sigma_2 &= 0.27 \\ \mathrm{cov}_{02} &= -0.14 \end{aligned}$$

$\chi^2$/NDF-fel szorozva az aszimptotikus hibák:

$$\begin{aligned} \tilde{\sigma}_0 &= 0.1000 \\ \tilde{\sigma}_1 &= 0.0454 \\ \tilde{\sigma}_2 &= 0.0384 \end{aligned}$$