Helymeghatározási alapelvek és módszerek

Helymeghatározás alapjai, 2. rész

Hollósi Gergely[1]

[1]Budapesti Műszaki és Gazdaságtudományi Egyetem
Távközlési és Médiainformatikai Tanszék

2015

A helymeghatározási feladat

- A helymeghatározás egyenlete ($\mathbf{l}$ a hely és $\mathbf{v}$ az orientáció):

$$\mathbf{x} = f(\mathbf{l}, \mathbf{v}) + \mathbf{x}_n$$

- Fennálló problémák:
  - az f függvény nem feltétlenül lineáris
  - a méréseket hibák és zajok terhelik
  - a több dimenziós egyenletrendszer túldefiniált lehet
- A fő kérés: *hogy oldjuk meg a helymeghatározási egyenletet?*
  - Természetesen függ a helybecslési technikától
  - Differenciálegyenlettel itt nem foglalkozunk

A helymeghatározási feladat

- A továbbiakban egy általános egyenlet megoldására törekszünk:

$$\mathbf{y} = f(\mathbf{x}) + \nu$$

- Az egyenletben ν-vel jelöljük a feltételezett zajt
- Az egyenlet lehetséges megoldási módszereivel lehetőség nyílik a helymeghatározási egyenlet megoldására
- Megoldási lehetőségek:
    - költségfüggvények $\rightarrow$ legkisebb négyzetes eltérések módszere
    - lineáris egyenletrendszerek megoldása
    - nemlineáris iteratív módszerek
    - valószínűségi becsléselmélet

Helymeghatározási feladat megoldása

Áttekintés

A helymeghatározási feladat megoldása

- Az egyenletrendszerek általában túldefiniáltak
  - megoldás ebben az esetben nem mindig létezik
- Keresni kell valamilyen „jóság" mértéket, amely szerint egy közelítő megoldás „jó" $\rightarrow$ *költségfüggvény*
- A $C(\mathbf{y}, \mathbf{x})$ költségfüggvény megadja, hogy egy választott megoldás költsége mekkora
- Minimalizáljuk a költséget, azaz keressük azt a paraméterhalmazt, amelyre
$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} C(\mathbf{y}, \mathbf{x})$$
- Mik a megfelelő költségfüggvények?

Költségfüggvény

A helymeghatározási feladat megoldása

- A leggyakoribb költségfüggvény az euklideszi normán alapul

$$C(\mathbf{y}, \mathbf{x}) = \|\mathbf{y} - f(\mathbf{x})\|^2$$

- Elterjedt neve: *legkisebb négyzetes eltérésen alapuló megoldás*
- Egyéb költségfüggvények is léteznek, ugyanakkor ritkán használják (pl. különbség abszolútértéke)

Költségfüggvény
A helymeghatározási feladat megoldása

- A leggyakoribb költségfüggvény az euklideszi normán alapul
$$C(\mathbf{y}, \mathbf{x}) = \|\mathbf{y} - f(\mathbf{x})\|^2$$
- Elterjedt neve: *legkisebb négyzetes eltérésen alapuló megoldás*
- Egyéb költségfüggvények is léteznek, ugyanakkor ritkán használják (pl. különbség abszolútértéke)

1. A négyzetreemelésnek a költségfüggvényben nincs elméleti jelentősége, azonban az euklideszi normában található gyökjel kiesik, így egyszerűbb számolni vele. Fontos megjegyezni ugyanakkor, hogy a négyzetreemelés ebben az esetben a minimumhelyen nem változtat.
2. Miért az euklideszi norma? Egyrészt azt a feltételezést tesszük, hogy a mérésekben csak zaj található, így tulajdonképpen – ahogy később látni fogjuk – a *maximum likelihood* becsléshez jutunk. Ezzel persze csak akkor érünk célt, ha $\mathbf{y}$ fizikailag értelmezhető mennyiség.

Lineáris egyenletek

A helymeghatározási feladat megoldása

- Gyakran előfordulnak lineáris egyenletek, tipikusan például vizuális helymeghatározás során
- Tegyük fel tehát, hogy

$$\mathbf{y} = f(\mathbf{x}) = A\mathbf{x} \qquad A \in \mathbb{C}^{m \times n}$$

- Amennyiben A négyzetes, és invertálható, $\mathbf{x} = A^{-1}\mathbf{y}$

Lineáris egyenletek - SVD felbontás (1)

A helymeghatározási feladat megoldása

- Minden $\mathbf{A}_{M \times N}$ mátrixra létezik egy ún. SVD (Singular Value Decomposition) felbontás

$$\mathbf{A}_{M \times N} = \mathbf{U}_{M \times M} \mathbf{\Sigma}_{M \times N} \mathbf{V}^H_{N \times N}$$

- $\mathbf{U}$ és $\mathbf{V}$ unitér (valós esetben ortogonális) mátrixok ($\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$ és $\mathbf{V}\mathbf{V}^H = \mathbf{V}^H\mathbf{V} = \mathbf{I}$)
- Σ mátrixban a főátlóban nemnegatív valós számok találhatók, máshol nulla
  - A főátló elemeit hívjuk szinguláris értékeknek (σ_n)

Lineáris egyenletek - SVD felbontás (1)

A helymeghatározási feladat megoldása

- Minden $\mathbf{A}_{M\times N}$ mátrixra létezik egy ún. SVD (Singular Value Decomposition) felbontás

$$\mathbf{A}_{M\times N} = \mathbf{U}_{M\times M}\mathbf{\Sigma}_{M\times N}\mathbf{V}^H_{N\times N}$$

- $\mathbf{U}$ és $\mathbf{V}$ unitér (valós esetben ortogonális) mátrixok ($\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$ és $\mathbf{V}\mathbf{V}^H = \mathbf{V}^H\mathbf{V} = \mathbf{I}$)
- Σ mátrixban a főátlóban nemnegatív valós számok találhatók, máshol nulla
  - A főátló elemeit hívjuk szinguláris értékeknek (σ_n)

1. A sajátértékekkel kapcsolatos hasonlóság, hogy az $\mathbf{A}$ szinguláris értékéhez létezik két olyan egységvektor, a baloldali szinguláris vektor u és a jobbolda szinguláris vektor v, amelyekre

$$\mathbf{A}v = \sigma u \qquad \mathbf{A}^H u = \sigma v$$

Lineáris egyenletek - SVD felbontás (2)

A helymeghatározási feladat megoldása

- Nem négyzetes esetben tételezzük fel, hogy $A = UDV^T$ az SVD felbontás szerint
- Minimalizáljuk a $\|A\mathbf{x} - \mathbf{y}\|$ kifejezést!

$$\|A\mathbf{x} - \mathbf{y}\| = \|UDV^T\mathbf{x} - \mathbf{y}\| = \|DV^T\mathbf{x} - U^T\mathbf{y}\|$$

- Legyen $\mathbf{y}' = U^T\mathbf{y}$ és $\mathbf{x}' = V^T\mathbf{x}$!

Lineáris egyenletek - SVD felbontás (3)

A helymeghatározási feladat megoldása

- Ekkor a $\|D\mathbf{x}' - \mathbf{y}'\|$ kifejezést kell minimalizálnunk:

$$\left[\begin{array}{cccc} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \\ \hline 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{array}\right] \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix} = \left[\begin{array}{c} y'_1 \\ y'_2 \\ \vdots \\ y'_n \\ \hline y'_{n+1} \\ \vdots \\ y'_m \end{array}\right] \tag{1}$$

- Nyilván a legközelebbi megoldás, ha $x'_i = y'_i / d_i$
- Végül, $\mathbf{x} = V\mathbf{x}'$

Lineáris egyenletek - Pszeudoinverz

A helymeghatározási feladat megoldása

- Másik lehetőség, hogy észrevesszük, hogy $A\mathbf{x}$ a legközelebbi pont $\mathbf{y}$-hoz.
- Ekkor $A\mathbf{x} - y$ merőleges A oszlopvektorai terére
- Ebből:

$$A^T(A\mathbf{x} - \mathbf{b}) = 0$$
$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{y}$$

- Ebből az $A^+ = (A^T A)^{-1} A^T$ kifejezés a pszeudoinverz

Lineáris egyenletek - Pszeudoinverz

A helymeghatározási feladat megoldása

- Másik lehetőség, hogy észrevesszük, hogy $A\mathbf{x}$ a legközelebbi pont $\mathbf{y}$-hoz.
- Ekkor $A\mathbf{x} - y$ merőleges A oszlopvektorai terére
- Ebből:

$$A^T(A\mathbf{x} - \mathbf{b}) = 0$$
$$\mathbf{x} = (A^TA)^{-1}A^T\mathbf{y}$$

- Ebből az $A^+ = (A^TA)^{-1}A^T$ kifejezés a pszeudoinverz

1. Az SVD alapú megoldás és a pszeudoinverz alapján könnyen belátható, hog $A^+ = VD^+U^T$, ahol D^+ $n \times m$ méretű diagonálmátrix a D mátrix átlóelemeinek reciprokaiból.

Lineáris egyenletek - Homogén lineáris egyenlet

A helymeghatározási feladat megoldása

- Speciális esetben homogén egyenletrendszer megoldása szükséges, azaz
$$A\mathbf{x} = 0$$
- Nyilván ha $\mathbf{x}$ megoldás, akkor $k\mathbf{x}$ is megoldás. Megkötés: $\|\mathbf{x}\| = 1$
- $\mathbf{x} = 0$ megoldás triviális és nem érdekes
- Legyen megint $A = UDV^T$ az SVD felbontás
$$\|UDV^T\mathbf{x}\| = \|DV^T\mathbf{x}\| \qquad \|\mathbf{x}\| = \|V^T\mathbf{x}\|$$
- Ekkor minimalizáljuk $\|D\mathbf{x}'\|$ kifejezést, ha $\|\mathbf{x}'\| = 1$ és $\mathbf{x}' = V^T\mathbf{x}$
- A megoldás $\mathbf{x}' = (0, 0, \cdots, 0, 1)^T$
- Végül tehát a megoldás a V mátrix utolsó oszlopa.

Nem lineáris egyenletrendszerek

Helymeghatározási egyenlet megoldása

- Zárt alakban csak az egyenletek és ismeretlenek számának egyezése esetén lehetséges
  - Ekkor is csak speciális esetekben
- Túldefiniált esetben költségfüggvény
- Közelítő iteratív módszerek léteznek $\rightarrow$ nem feltétlen konvergens
- Gyakorlati problémák esetén azonban legtöbbször konvergens és hatékony
- Probléma, hogy kell egy jó kiindulási pont
  - Például az egyenletrendszer töredékét megoldjuk zárt alakban

Nem lineáris egyenletrendszerek

Helymeghatározási egyenlet megoldása

- Legyen egy $g(\mathbf{x})$ skalár értékű függvényünk, tehát $g : \mathbb{R}^n \to \mathbb{R}$
- Keressük meg a g minimumát $\mathbf{x}$ függvényében!
- Taylor-sor $\mathbf{x}_0$ körül

$$g(\mathbf{x_0} + \Delta) = g(\mathbf{x_0}) + g_{\mathbf{x}}(\mathbf{x_0})\Delta + \Delta^T g_{\mathbf{xx}}(\mathbf{x_0})\Delta/2 + \dots$$

- Az alsóindex $\mathbf{x}$ a deriválást jelenti (gradiens és Hesse-mátrix)
- Az első három tagot tekintsük, és keressük a minimumot $\to$ deriválás

$$g_{\mathbf{xx}}\Delta = -g_{\mathbf{x}}$$

- A módszer tehát Δ irányba lépegetve megtalálni a minimumot
  1. Oldjuk meg a $g_{\mathbf{xx}}(\mathbf{x_n})\Delta = -g_{\mathbf{x}}(\mathbf{x}_n)$
  2. Legyen $\mathbf{x}_{n+1} = \mathbf{x}_n + \Delta$

Nem lineáris egyenletrendszerek

Helymeghatározási egyenlet megoldása

- Hogyan alkalmazzuk ezt a legkisebb négyzetes eltérésre?
- Legyen $\epsilon = f(\mathbf{x}) - \mathbf{y}$, s használjuk a költségfüggvényt

$$g(\mathbf{x}) \triangleq \frac{1}{2} C(\mathbf{y}, \mathbf{x}) = \frac{1}{2} \|\epsilon(\mathbf{x})\|^2 = \epsilon(\mathbf{x})^T \epsilon(\mathbf{x})/2$$

- Vegyük észre, hogy $g_{\mathbf{x}} = \epsilon_{\mathbf{x}}^T \epsilon$, amelyben $\epsilon_{\mathbf{x}} = f_{\mathbf{x}} = J$ az f Jacobi-mátrixsza
- A Hesse-mátrix: $g_{\mathbf{xx}} = \epsilon_{\mathbf{x}}^T \epsilon_{\mathbf{x}} + \epsilon_{\mathbf{xx}}^T \epsilon$
- Ezek alapján négy alapvető módszert különböztetünk meg:
  1. Newton-módszer
  2. Gauss-Newton módszer
  3. gradiens módszer
  4. Levenberg-Marquardt módszer

Nem lineáris egyenletrendszerek – Newton-módszer

Helymeghatározási egyenlet megoldása

- A frissítési egyenlete

$$g_{\mathbf{xx}}\Delta = -g_{\mathbf{x}}$$

- Problémás, hogy a Hesse-mátrixra is szükség van
- Gyors konvergencia, nagy számításigény

Nem lineáris egyenletrendszerek – Gauss-Newton-módszer

Helymeghatározási egyenlet megoldása

- Közelítsük a Hesse mátrixot:

$$g_{\mathbf{xx}} = \epsilon_{\mathbf{x}}^T \epsilon_{\mathbf{x}} + \epsilon_{\mathbf{xx}}^T \epsilon = \epsilon_{\mathbf{x}}^T \epsilon_{\mathbf{x}} = J^T J$$

- A frissítési egyenlete

$$J^T J \Delta = -J^T \epsilon$$

- Tehát $\mathbf{x}_{n+1} = \mathbf{x}_n - (J^T J)^{-1} J^T (f(\mathbf{x}) - \mathbf{y})$

- Nincs szükség a Hesse-mátrixra, a másodfokú deriváltakra
- Gyors konvergencia, kisebb számításigény

Nem lineáris egyenletrendszerek – gradiens-módszer

Helymeghatározási egyenlet megoldása

- Egyszerűen a gradiens irányába halad
- A frissítési egyenlete

$$\lambda\Delta = -g_{\mathbf{x}}$$

- Nincs szükség a Hesse-mátrixra, a másodfokú deriváltakra
- Lassú konvergencia, hajlamos oszcillációra

Nem lineáris egyenletrendszerek – Levenberg-Marquardt-módszer

Helymeghatározási egyenlet megoldása

- A Gauss-Newton és a gradiens módszer keverése
- A frissítési egyenlete

$$(J^T J + \lambda I)\Delta = -J^T \epsilon$$

- Működése
  1. Kell egy kezdeti λ érték
  2. Megoldjuk a frissítési egyenletet egyenletet
  3. Ha ϵ növekedett, akkor $\lambda_{n+1} = R\lambda_n$, R tipikusan 10, és újra megoldjuk a frissítési egyenletet (2)
  4. Ha ϵ csökkent, akkor $\lambda_{n+1} = \dfrac{\lambda_n}{R}$, ahol R tipikusan 10
  5. Frissítjük a paramétereket, $\mathbf{x}_{n+1} = \mathbf{x}_n + \Delta$
  6. Újabb iteráció (2)
- Gyorsabb konvergencia, de stabilabb megoldás
- Vizuális módszereknél gyakran használják

Valószínűségi becslés – ismétlés

Helymeghatározási egyenlet megoldása

- Valószínűségi változó
- Valószínűségi eloszlásfüggvény és sűrűségfüggvény
- Feltételes valószínűség
- Várható érték és szórás
- Kovariancia és korreláció, kovariancia mátrix
- Nevezetes eloszlások
  - többváltozós normális-eloszlás
- Centrális határeloszlástétel

Valószínűségi becslés – ismétlés
Helymeghatározási egyenlet megoldása

- Valószínűségi változó
- Valószínűségi eloszlásfüggvény és sűrűségfüggvény
- Feltételes valószínűség
- Várható érték és szórás
- Kovariancia és korreláció, kovariancia mátrix
- Nevezetes eloszlások
  - többváltozós normális-eloszlás
- Centrális határeloszlástétel

1. $P(a > \mathbb{X}) = \int_{-\infty}^{a} p(x)dx$
2. $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$
3. Várható érték: $E(X) = \int_{-\infty}^{+\infty} xp(x)\,dx \quad E(X) = \sum_{i=1}^{\infty} x_i p_i$
4. Szórás: $D^2(X) = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right]$
5. Becslés mintákból: $\hat{\sigma} = \frac{1}{n-1}\sum_i (x_i - \overline{x})^2$
6. Kovariancia: $\sigma(X, Y) = \mathrm{E}\left[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])\right]$
7. Becslés: $\hat{\sigma}(X, Y) = \frac{1}{n-1} M_x M_y$
8. Bernoulli eloszlás: $p(k; \mu) = \mu^k (1-\mu)^{1-k} \quad k \in 0, 1$
9. Normális eloszlás:
$$p_{\mathbf{x}}(x_1, \ldots, x_k) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$
10. Centrális határeloszlás: N darab azonos eloszlású és független valószínűségi változó összege normális eloszláshoz tart, amennyiben $N \to \infty$

Becsléselmélet – Estimation Theory

Helymeghatározási egyenlet megoldása

- Tapasztalati megfigyelések alapján paraméterek becslése
- $\mathbf{x}$ mérések, θ paraméterhalmaz, $p(\mathbf{x}|\theta)$ modell
- Becslő a megfigyelések alapján $\hat{\theta}(\mathbf{x})$, az i-dik paraméter becslője $\hat{\theta}_i(\mathbf{x})$
- Cramer-Rao tétel az elméleti minimum a bizonytalanságra

$$\mathrm{var}\left[\theta_i - \hat{\theta}_i(\mathbf{x})\right] \geq -E\left[\frac{\partial^2 \ln p(\mathbf{x}|\theta)}{\partial^2 \theta_i}\right]$$

- Például egy egyszerű σ szórású Gauss-i folyamat
  $x_i = \theta + n_i \quad n = 1, \ldots, N$

$$\mathrm{var}\left[\theta - \hat{\theta}(\mathbf{x})\right] = \frac{\sigma^2}{N}$$

Valószínűségi becslés – áttekintés

Helymeghatározási egyenlet megoldása

- Fogalmazzuk meg a becslés feladatát!
- $z_1, z_2, \ldots, z_N$ méréshalmaz
- Paraméterbecslés: $z_{1:N} = \{z_1, z_2, \ldots, z_N\} \to \theta$
  - folyamatos
  - kötegelt
- Állapotbecslés $z_{1:N} = \{z_1, z_2, \ldots, z_N\} \to x_{1:N} = \{x_1, x_2, \ldots, x_N\}$
  - simítás
  - szűrés
  - extrapoláció

Valószínűségi becslés – Maximum Likelihood

Helymeghatározási egyenlet megoldása

- Tekintsük az együttes eloszlást: $p(z_{1:N}|\theta)$
- Keressük azt a θ paramétert, amelyre a mérési sorozat a legvalószínűbb
- Likelihood függvény: $\mathcal{L}(\theta; z_{1:N}) = p(z_{1:N}|\theta)$
  - θ szabadon választható, tehát $\mathcal{L}$ a θ függvénye
- A Maximum Likelihood módszer elve tehát:

$$\theta_{ML} = \arg\max_{\theta} \mathcal{L}(\theta; z_{1:N})$$

Valószínűségi becslés – Maximum Likelihood (1. példa)

Helymeghatározási egyenlet megoldása

- Nézzünk egy – akár cinkelt – pénzérmét
- Legyen X valószínűségi változó 1, ha fejre esik, 0, ha írásra
- $P(X = 1) = \mu$ és $P(X = 0) = 1 - \mu$, Bernoulli-eloszlás
- Egy $z_0, z_1, \ldots, z_N$ dobássorozatra becsüljük meg a μ értékét, ha az egyes dobások függetlenek

$$\mathcal{L}(\mu) = p(z_{1:N}|\mu) = \prod_{i=1}^{N} \mu^{z_i}(1-\mu)^{1-z_i}$$

- A megoldás:

$$\mu_{ML} = \sum_{i=1}^{N} z_i$$

Valószínűségi becslés – Maximum Likelihood (1. példa)

Helymeghatározási egyenlet megoldása

- Nézzünk egy – akár cinkelt – pénzérmét
- Legyen X valószínűségi változó 1, ha fejre esik, 0, ha írásra
- $P(X=1)=\mu$ és $P(X=0)=1-\mu$, Bernoulli-eloszlás
- Egy $z_0, z_1, \ldots, z_N$ dobássorozatra becsüljük meg a μ értékét, ha az egyes dobások függetlenek

$$\mathcal{L}(\mu) = p(z_{1:N}|\mu) = \prod_{i=1}^{N} \mu^{z_i}(1-\mu)^{1-z_i}$$

- A megoldás:

$$\mu_{ML} = \sum_{i=1}^{N} z_i$$

$$\begin{aligned}\ln \mathcal{L} &= \sum_{i=1}^{N} \ln \mu^{z_i}(1-\mu)^{1-z_i} = \ln \mu \sum_{i=1}^{N} z_i + \ln(1-\mu)\sum_{i=1}^{N}(1-z_i) \\ &= N\bar{z}\ln\mu + (N - N\bar{z})\ln(1-\mu)\end{aligned}$$

Deriválás μ szerint, és maximumkeresés:

$$N\bar{z}\frac{1}{\mu} - (N - N\bar{z})\frac{1}{1-\mu} = 0$$

$$N\bar{z} - N\bar{z}\mu - N\mu + N\bar{z}\mu = 0$$

$$\mu = \bar{z} = \frac{1}{N}\sum_{i=1}^{N} z_i$$

Valószínűségi becslés – Maximum Likelihood (2. példa)

Helymeghatározási egyenlet megoldása

- Nézzünk egy $z_i = f(x)$ függvényt, ahol az z mérések normális eloszlású zajjal terheltek, és a mérések függetlenek, de azonos eloszlásúak!
- A likelihood függvény:

$$\mathcal{L}(x) = \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(z_i - f(x))^2}{2\sigma^2}\right)$$

- Levezetés logaritmus függvénnyel, az eredmény:

$$\arg\max_x \mathcal{L} = \arg\min_x \sum_{i=1}^{N} (z_i - f(x))^2$$

- Legkisebb négyzetes eltérés $\rightarrow$ ML becslés normális eloszlású zajban

Valószínűségi becslés – Maximum Likelihood (2. példa)

Helymeghatározási egyenlet megoldása

- Nézzünk egy $z_i = f(x)$ függvényt, ahol az z mérések normális eloszlású zajjal terheltek, és a mérések függetlenek, de azonos eloszlásúak!
- A likelihood függvény:

$$\mathcal{L}(x) = \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(z_i - f(x))^2}{2\sigma^2}\right)$$

- Levezetés logaritmus függvénnyel, az eredmény:

$$\arg\max_x \mathcal{L} = \arg\min_x \sum_{i=1}^{N} (z_i - f(x))^2$$

- Legkisebb négyzetes eltérés → ML becslés normális eloszlású zajban

A megoldás az eloszlásfüggvény logaritmusának kiszámításával egyből adódik:

$$\ln \mathcal{L} = \sum_{i=1}^{N} \ln \frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2\sigma^2} \sum_{i=1}^{N} (z_i - f(x))^2$$

Valószínűségi becslés – Maximum Likelihood

Helymeghatározási egyenlet megoldása

- Egyszerű és hatékony keretek valószínűségi becsléshez
  - Bonyolult problémáknál bonyolult egyenletrendszerek
- Nem minden esetben helyes
  - Tekintsük a pénzfeldobási feladatot a 1,1,1 dobási sorozatra!
  - $\mu = 1$, ami nem feltétlenül értelmes
  - Hiányzik az előzetes ismeret

Valószínűségi becslés – Bayes becslés

Helymeghatározási egyenlet megoldása

- Keressük meg a ,,legvalószínűbb" megoldást!
- Matematikailag, számítsuk ki a $p(\theta|z_{1:N})$ valószínűséget!
  - állapotbecslés esetén a $p(x_{1:N}|z_{1:N})$ valószínűséget
  - Az eloszlásból már készíthetünk pontbecslést, pl. az eloszlás maximumánál
- Lényegében tehát magát a meghatározandó mennyiséget is valószínűségi változónak kezeljük $\rightarrow$ valódi valószínűségi becslés

Valószínűségi becslés – Bayes tétel

Helymeghatározási egyenlet megoldása

- A Bayes becslés a Bayes tételen alapul. Legyen egy paraméter θ és egy mérés z!
- Ekkor

Hipotézis **Prior**

$$\textbf{Posterior} \longrightarrow \quad p(\theta|z) = \frac{p(z|\theta)p(\theta)}{p(z)}$$

- A hipotézis egyfajta modellismeret
- A prior a paraméterrel kapcsolatos előismereteink és feltételezéseink („a priori" ismeretek)
- A posterior a paraméterrel kapcsolatos ismereteink a mérések ismeretében („a posteriori")

Valószínűségi becslés – MAP

Helymeghatározási egyenlet megoldása

- A posterior eloszlás előállítása általában bonyolult
  - A $p(z) = \int_\vartheta p(z|\vartheta)p(\vartheta)\,\mathrm{d}\vartheta$ nem ismert, és integrálással számítható
- Bizonyos eloszlásokra könnyebben számítható, lásd. később

- A becslés során egy posterior eloszláshoz jutunk – mi a helyzet, ha egy konkrét értékre vagyunk kíváncsiak?
- Maximum a Posteriori (MAP) becslés

$$\theta_{MAP} = \arg\max_\theta p(\theta|z) \propto \arg\max_\theta p(z|\theta)p(\theta)$$

- Amennyiben a prior informális, akkor maximum likelihood becslés
  - a prior nem tartalmaz információt, nincs előismeretünk

Rekurzív Bayes szűrés

Helymeghatározási egyenlet megoldása

- Kiemelkedő mérnöki fontosságú
- A feladat $p(x_k|z_{1:k})$ meghatározása minden k mérésre rekurzívan
- Specifikus modell: Markov-folyamat

$$
\begin{array}{ccccccc}
\dashrightarrow & x_{k-1} & \longrightarrow & x_k & \longrightarrow & x_{k+1} & \dashrightarrow \\
& \downarrow & & \downarrow & & \downarrow & \\
& z_{k-1} & & z_k & & z_{k+1} &
\end{array}
$$

- Az állapottér modell tehát:

$$
\begin{aligned}
x_0 &\sim p(x_0) && \rightarrow \text{prior információk} \\
x_k &\sim p(x_k|x_{k-1}) && \rightarrow \text{dinamikus modell} \\
z_k &\sim p(z_k|x_k) && \rightarrow \text{mérési modell}
\end{aligned}
$$

Rekurzív Bayes szűrés – általános megoldás

Helymeghatározási egyenlet megoldása

- Írjuk fel az együttes eloszlást!

$$
\begin{aligned}
p(x_k, x_{k-1}|z_{1:k-1}) &= p(x_k|x_{k-1}, z_{1:k-1})p(x_{k-1}|z_{1:k-1}) = \\
&= p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})
\end{aligned}
$$

- A **predikciós** lépés

$$
p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})\mathrm{d}x_{k-1}
$$

- A **frissítési** lépés

$$
\begin{aligned}
p(x_k|z_{1:k}) &= \frac{1}{Z_k}p(z_k|x_k, z_{1:k-1})p(x_k|z_{1:k-1}) \\
&= \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{\int p(z_k|x_k)p(x_k|z_{1:k-1})\mathrm{d}x_k}
\end{aligned}
$$

Rekurzív Bayes szűrés – általános megoldás

Helymeghatározási egyenlet megoldása

- Írjuk fel az együttes eloszlást!

$$\begin{aligned} p(x_k, x_{k-1}|z_{1:k-1}) &= p(x_k|x_{k-1}, z_{1:k-1})p(x_{k-1}|z_{1:k-1}) = \\ &= p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1}) \end{aligned}$$

- A **predikciós** lépés

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})\mathrm{d}x_{k-1}$$

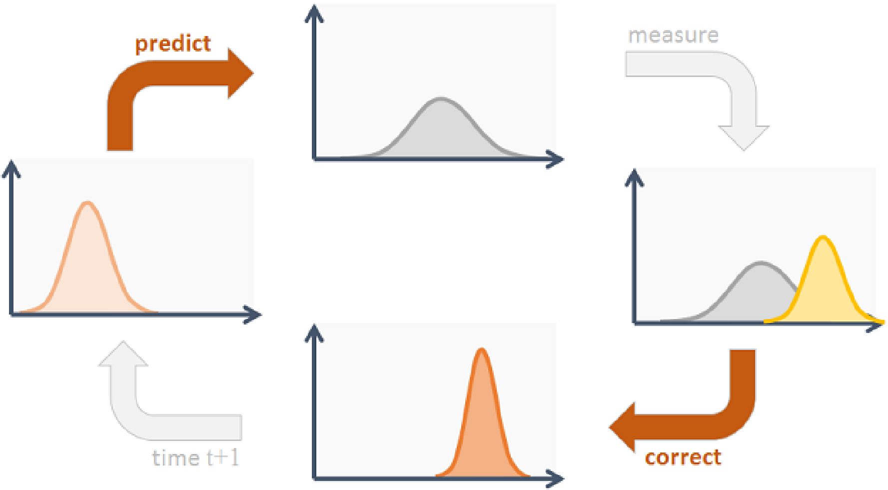
- A **frissítési** lépés

$$\begin{aligned} p(x_k|z_{1:k}) &= \frac{1}{Z_k} p(z_k|x_k, z_{1:k-1})p(x_k|z_{1:k-1}) \\ &= \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{\int p(z_k|x_k)p(x_k|z_{1:k-1})\mathrm{d}x_k} \end{aligned}$$

1. Az első egyenletben egyszerűen a feltételes valószínűség definícióját, majd a Markov-folyamat definícióját alkalmaztuk.
2. A második lépés az ún. Chapman-Kolmogorov egyenlőség, a teljes eseménytér összegzése egy valószínűségi változó szerint.
3. A harmadik egyenlet egyszerűen a Bayes-tétel felírása.

Rekurzív Bayes szűrés – általános megoldás

Helymeghatározási egyenlet megoldása



Rekurzív Bayes szűrés – speciális megoldás

Helymeghatározási egyenlet megoldása

- Az általános megoldás bonyolult $\rightarrow$ megkötések esetén egyszerűbb
- Speciális megoldási lehetőségek:
  - Kalman-szűrő
  - kiterjesztett Kalman-szűrő
  - részecske szűrő
- Egyéb, ritkábban használt megoldások
  - unscented Kalman-szűrő
  - Gauss-szűrő
  - Rauch-Tung-Striebel simító
  - stb.

Rekurzív Bayes szűrés – Kalman-szűrő

Helymeghatározási egyenlet megoldása

- Lineáris modell, normális eloszlás

$$x_k = A_{k-1}x_{k-1} + q_{k-1} \qquad q_{k-1} \sim N(0, Q_{k-1})$$
$$z_k = H_k x_k + r_k \qquad r_k \sim N(0, R_k)$$

- Valószínűségek szerint

$$p(x_k|x_{k-1}) = N(x_k|A_{k-1}x_{k-1}, Q_{k-1})$$
$$p(z_k|x_k) = N(z_k|H_k x_k, R_k)$$

- Ebben az esetben x_k eloszlása is normális:

$$p(x_k|z_{1:k}) \sim N(x_k|m_k, P_k)$$

Rekurzív Bayes szűrés – Kalman-szűrő

Helymeghatározási egyenlet megoldása

- A **predikciós** lépés:

$$m_k^- = A_{k-1} m_{k-1}$$
$$P_k^- = A_{k-1} P_{k-1} A_{k-1}^T + Q_{k-1}$$

- A **frissítési** lépés:

$$v_k = z_k - H_k m_k^-$$
$$S_k = H_k P_k^- H_k^T + R_k$$
$$K_k = P_k^- H_k^T S_k^{-1}$$
$$m_k = m_k^- + K_k v_k$$
$$P_k = P_k^- - K_k S_k K_k^T$$

Rekurzív Bayes szűrés – Kalman-szűrő példa
Helymeghatározási egyenlet megoldása

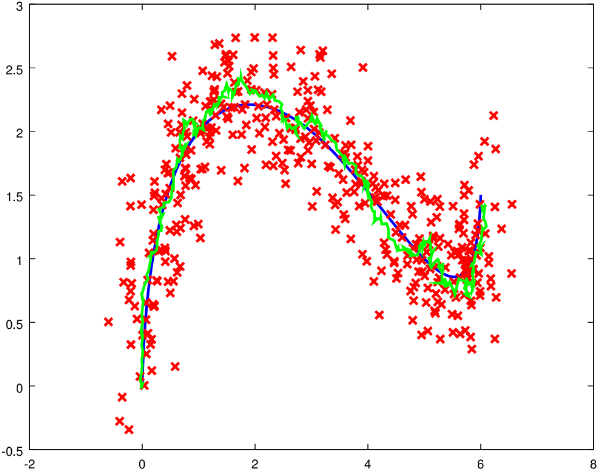
- Egy $l(l_x, l_y)$ helyzetű és $v(v_x, v_y)$ sebességű autó modellje.
- Az állapotvektor $x = [l_x, l_y, v_x, v_y]^T$

$$A = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Rekurzív Bayes szűrés – Kalman-szűrő példa

Helymeghatározási egyenlet megoldása



Rekurzív Bayes szűrés – Kiterjesztett Kalman-szűrő

Helymeghatározási egyenlet megoldása

- Nem lineáris modell, normális eloszlás
- A modell linearizálása a várható érték körül
- Állapotegyenletek:

$$\begin{aligned} x_k &= f(x_{k-1}) + q_{k-1} \qquad & q_{k-1} &\sim N(0, Q_{k-1}) \\ z_k &= h(x_k) + r_k \qquad & r_k &\sim N(0, R_k) \end{aligned}$$

- x_k eloszlását is normális eloszlással közelítjük:

$$p(x_k | z_{1:k}) \sim N(x_k | m_k, P_k)$$

Rekurzív Bayes szűrés – Kiterjesztett Kalman-szűrő

Helymeghatározási egyenlet megoldása

- A **predikciós** lépés:

$$\begin{aligned} m_k^- &= f(m_{k-1}) \\ P_k^- &= F_x(m_{k-1}) P_{k-1} F_x^T(m_{k-1}) + Q_{k-1} \end{aligned}$$

- A **frissítési** lépés:

$$\begin{aligned} v_k &= z_k - h(m_k^-) \\ S_k &= H_x(m_k^-) P_k^- H_x^T(m_k^-) + R_k \\ K_k &= P_k^- H_x^T(m_k^-) S_k^{-1} \\ m_k &= m_k^- + K_k v_k \\ P_k &= P_k^- - K_k S_k K_k^T \end{aligned}$$

- $F_x(m_k)$ és $H_x(m_k)$ az f és h függvény Jacobi-mátrixsza x szerint az m_k helyen

Rekurzív Bayes szűrés – Kiterjesztett Kalman-szűrő példa

Helymeghatározási egyenlet megoldása

- Szögmérésből származó információk alapján történő helymeghatározás

$$\mathbf{l}_k = f(\mathbf{l}_{k-1}) = \mathbf{l}_{k-1} + q$$

$$\alpha_n = h_n(\mathbf{l}_k) = \cos^{-1} \frac{(\mathbf{l} - \mathbf{p}_n)\mathbf{r}_n}{|\ \mathbf{l} - \mathbf{p}_n\ |} + r$$

Rekurzív Bayes szűrés – részecske-szűrő

Helymeghatározási egyenlet megoldása

- Tetszőleges modellre és eloszlásra
- Tulajdonképpen numerikus – valószínűségi – integrálást csinálunk
- Lényege a Monte-Carlo integrálás

$$\int_{\Omega} f(x)\mathrm{d}x \approx V \frac{1}{N} \sum_{i=1}^{N} f(x^i) \qquad V = \int_{\Omega} \mathrm{d}x$$

- A konvergencia gyorsítható
  - fontosság szerinti mintavétel (közel azonos eloszlásból)

- A módszer előnye, hogy tetszőleges esetben működik
- Hátránya, hogy nagyon számításigényes

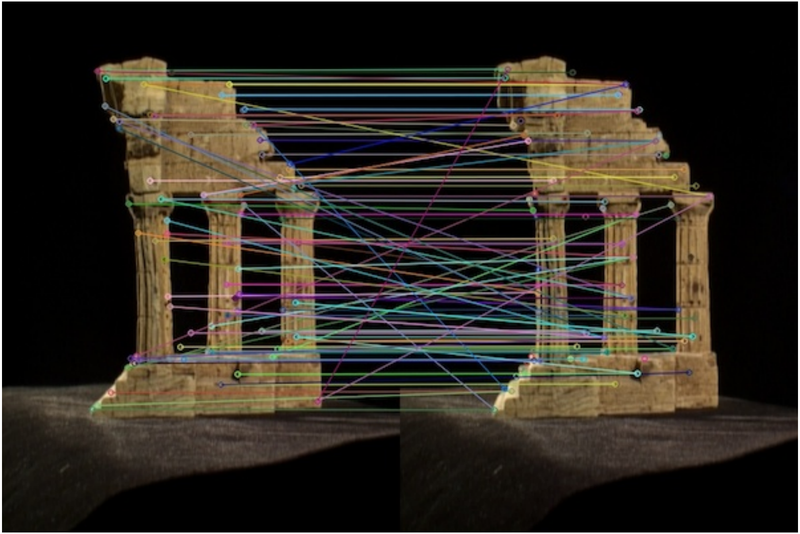
Robusztus megoldások

Helymeghatározási egyenlet megoldása

- Bizonyos esetekben a hibákat nem tudjuk modellezni
  - Kevés információ
  - Túl bonyolult
- A mérések egy része megfelel a modellnek $\rightarrow$ inlier
- A mérések másik része teljesen véletlenszerű $\rightarrow$ outlier
- Az outlier mérésekre a modellünk nem érvényes, a belőle származó becslés pontatlan
- A feladat az inlier mérések megkeresése

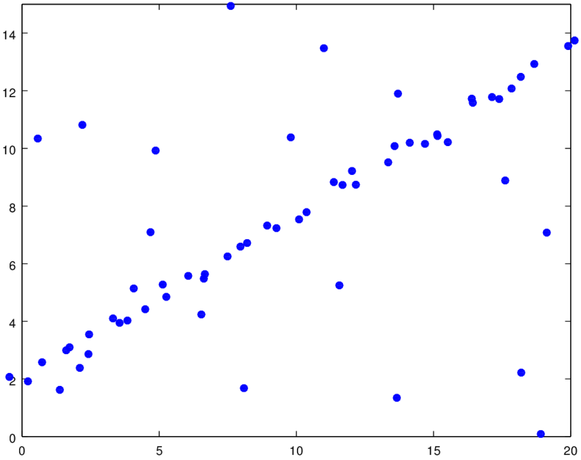
RANSAC

Helymeghatározási egyenlet megoldása



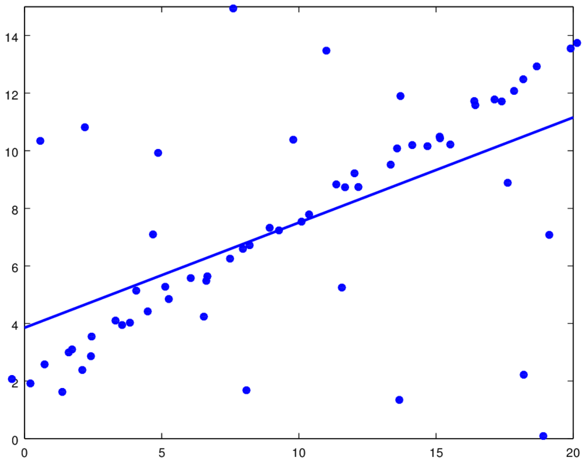
RANSAC

Helymeghatározási egyenlet megoldása



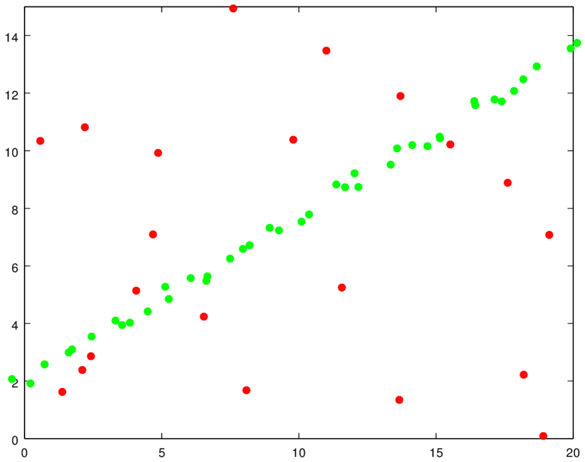
RANSAC

Helymeghatározási egyenlet megoldása



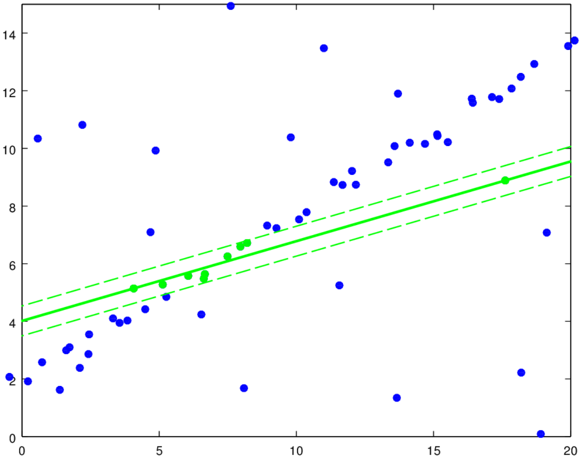
RANSAC

Helymeghatározási egyenlet megoldása



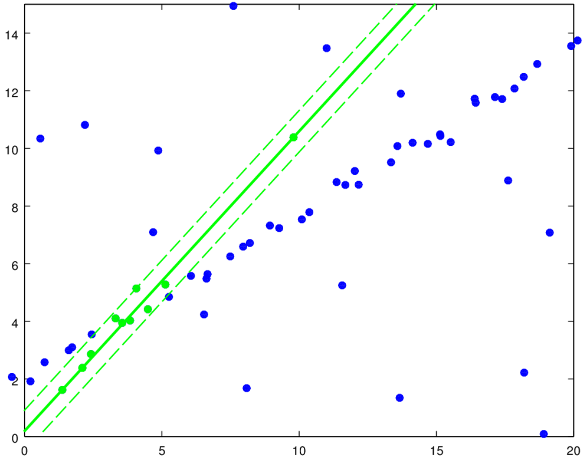
RANSAC

Helymeghatározási egyenlet megoldása



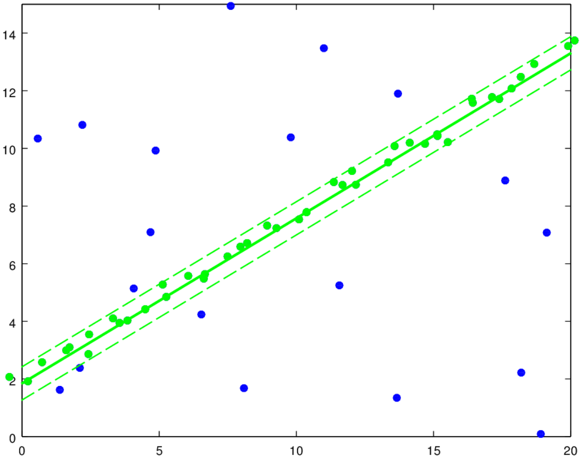
RANSAC

Helymeghatározási egyenlet megoldása



RANSAC

Helymeghatározási egyenlet megoldása



RANSAC

Helymeghatározási egyenlet megoldása

- RANSAC – Random Sample Consensus
- Nem determinisztikus módszer egy adott modellnek megfelelő mérések megtalálására

- Legyen egy $M(\theta)$ modellünk, ahol θ a modell paraméterei
  - Például az egyenesnél $\theta = \{m, b\}$ (ugyanis $y = mx + b$)
- Ne az összes mérést használjuk, hanem a legkisebb részhalmazát, ami egyértelműen meghatározza θ értékét
- Határozzuk meg egy valamilyen ϵ szórásnak megfelelő mérések számát $\rightarrow$ inlier-ek
- Ismételjük, majd válasszuk azt a θ-t, amely a legtöbb inlier-t adja

RANSAC

Helymeghatározási egyenlet megoldása

- Többféle megállási feltétel
  - inlier-ek száma
  - az inlier-ek hibájának szórása
  - valószínűségi feltétel
- Futásidő nagy lehet
  - θ meghatározása k mérésből, és N az összes mérés száma, akkor $\binom{N}{k}$ kombináció
  - Például az előző példa (60 db pont, min. 2 pont): 1770 próbálkozás
  - Párhuzamosítható
- Az ϵ hibahatár megállapítása önkényes, tapasztalati alapon vagy várakozásoknak megfelelően

RANSAC – valószínűségi megállási feltétel

Helymeghatározási egyenlet megoldása

- p valószínűséggel találjunk meg jó megoldást
- Minimum s darab minta a modellparaméterek meghatározásához
- w a valószínűsége, hogy egy minta inlier ($\epsilon = 1 - w$, hogy outlier)
- Ekkor $(1 - w^s)^N = 1 - p$, ha N ciklust futtatunk

$$N = \frac{\log(1-p)}{\log(1-(1-\epsilon)^s)}$$

	Outlier valószínűség (ϵ)						
Minták (s)	5%	10%	20%	25%	30%	40%	50%
2	2	3	5	6	7	11	17
3	3	4	7	9	11	19	35
4	3	5	9	13	17	34	72
5	4	6	12	17	26	57	146
6	4	7	16	24	37	97	293

RANSAC – valószínűségi megállási feltétel

Helymeghatározási egyenlet megoldása

- Az outlier-ek valószínűsége, ϵ adaptívan is becsülhető
- Ha találunk egy megoldást, amelyre N_{inlier} darab inlier található, akkor legalább ennyi inlier található

- $N = \infty$, sample_count=0
- Amíg $N >$ sample_count
  - Válasszunk egy mintahalmazt, és határozzuk meg az inlier-ek számát
  - Legyen $\epsilon = 1 - \frac{N_{\text{inlier}}}{\text{minták száma}}$
  - Határozzuk meg N értékét (lásd. előbb)
  - Növeljük eggyel a sample_count értékét
- Vége

Least Median Squares Estimation – LMedS

Helymeghatározási egyenlet megoldása

- A RANSAC robusztus technika, de kell egy hibahatárérték
  - Tehát ismerni kell a modellen belüli hiba szórását
- A medián is használható
  - A négyzetösszeget nagyon torzítja az outlier
  - Legkisebb abszolút eltérések összege
- Minimalizáljuk a hiba mediánját:

$$\hat{\theta} = \arg\min_{\theta} \operatorname{med}_{i} r_i^2(\theta, m)$$

Least Median Squares Estimation – LMedS

Helymeghatározási egyenlet megoldása

- Zárt alakban nem megoldható
- Monte Carlo megoldás $\rightarrow$ mintavétel
  - Válasszunk ki minimális mennyiségű pontot
  - Határozzuk meg a modell paramétereket
  - Számítsuk ki a hiba mediánját a paraméterekhez képest
  - A megoldás a legkisebb mediánnal rendelkező minta
- Hasonló a RANSAC-hez, de nem kell hibahatár
- Ciklusok száma a RANSAC-hez hasonlóan számítható

Least Median Squares Estimation – LMedS

Helymeghatározási egyenlet megoldása

- Előnyök
  - Nem kell a hibahatárt megadni, ismeretlen hibára is jól működik
  - Párhuzamosítható, egyszerű
- Hátrány
  - Normális eloszlású zajra nem ideális
  - Az outlier-ek száma nem lehet több, mint 50 %