Optimalno filtriranje

Iščemo optimalen predpis za optimizacijo realnega sistema S na modelski sistem M, če gledamo v realnem sistemu spremenljivko X_S v modelskem pa jo ponazorimo z X_M.

Zahteve:

1. Šibka sklopitev med S in $M (čim manj vpliva)
2. X_M mora biti berljiva količina (odvisna od t)
3. Moramo imeti oceno stopnje usklajenosti

$\lim_{t\to\infty}\langle (X_M - X_S)^2\rangle = \ldots$

1. Dinamika (to so linearne diferencialne enačbe) za X_S in X_S **naj bo kar se da** podobna.

Optimalno združevanje

Imejmo dve ločeni opazovanji spremenljivke x. Izmerek označimo s z. Poleg prave vrednosti ima ta prištet še nek merilni šum:

$$z = x + r$$

Zanima nas po čem kakšni porazdelitvi je porazdeljen merilni šum. Izkaže se, da je po **Gaussovi porazdelitvi**:

$$\frac{dP}{dr} = \mathcal{N}(0, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{r^2}{2\sigma^2}}$$

Če poračunamo pričakovane vrednost $\langle r\rangle$ dobimo:

$$\langle r\rangle = \int_{-\infty}^{\infty} \frac{dP}{dr} r\, dr = 0$$

zaradi sodosti Gaussove porazdelitve in lihosti linearne funkcije. Za $\langle r^2\rangle$ pa dobimo:

$$\langle r^2\rangle \neq 0$$

$$\langle r^2\rangle = \int \frac{dP}{dr} r^2\, dr = \frac{1}{\sqrt{2\pi}\sigma}\int e^{-r^2/2\sigma^2} r^2\, dr =$$

Tu uvedemo novo spremenljivko $u = r^2/2\sigma^2$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int e^{-u^2} 2\sigma^3 u^2 \sqrt{2}\, du = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-u^2} u^2\, du\, (2\sigma^2) =$$

$$= \sigma^2$$

Tako lahko naši dve opazovanji spremenljivke x zapišemo kot:

$$\bar{z}_1,\ \sigma_1$$
$$\bar{z}_2,\ \sigma_2$$

Vsota **neodvisnih** naključnih spremenljivk teži k normalni porazdelitvi (spomni se CLT)

[Zgled: Brumov šum]

Kot zgled si poglejmo Brumov šum (oz. včasih znan kar kot Brum). To so motnje zaradi napajanja z AC napetostjo. Torej lahko opišemo napetost nekako takole:

$$U = U_0 \cos\omega t$$

Verjetnost, da ob nekem času izvedemo meritev je:

$$\frac{dP}{dt} = \frac{1}{T/2}$$

in je na intervalu $[0, T/2]$ konstantna. Verjetnost glede na napetost pa ni konstantna $\frac{dP}{dU} \neq \mathrm{konst.}$ Njen odvod se glasi:

$$dU = -U_0\omega \sin\omega t \, dt$$

in če naredim potem posredni odvod, lahko zapišemo:

$$\frac{dP}{dt}\left(\left|\frac{dt}{dU}\right|\right) = \frac{dP}{dt}\left|\frac{1}{-U_0\omega\sqrt{\sin^2\omega t}}\right| =$$

$$= \frac{dP}{dt}\frac{1}{U_0\omega\sqrt{1-\cos^2\omega t}} = \frac{dP}{dt}\frac{T}{2\pi\sqrt{U_0^2 - U^2}} =$$

$$= \frac{1}{T/2}\frac{1}{2\pi}\frac{T}{\sqrt{U_0^2 - U^2}} = \frac{1}{\pi}\frac{1}{\sqrt{U_0^2 - U^2}}$$

Tu si prosim oglejte slike v zvezku saj so precej pomembne za razumevanje, ampak in short, oblika še zdaleč ni Gaussovka. K sreči nas reši **Centralni Limitni Teorem**. Torej če imamo več prispevkov po brumu gredo proti Gaussu. To se dobro pozna že pri recimo $N = 10$.

Povprečevanje

Recimo, da imamo N meritev, ki so porazdeljene normalno okoli spremenljivke x z varianco σ:

$$\frac{dP}{dz_i} = \mathcal{N}(x, \sigma)$$

Povprečje lahko zapišemo potem kot:

$$\bar{z} = \frac{1}{N}\sum z_i$$

Za pričakovano vrednost merilnega šuma vemo:

$$\overline{(z_i - x)} = \bar{r}_i = 0 \quad \Rightarrow \quad \overline{(\bar{z} - x)} = 0$$

Izračunajmo še drugi moment:

$$\displaylines{ \overline{(z_i - x^2)} = \frac{1}{N^2}\overline{\left(\sum\left(z_i -x\right)\right)^2} = \[3mm]

=\frac{1}{N^2}\left[\overline{\sum(z_i - x)^2} + \overline{\sum_{i,j}{(z_i - x)(z_j - x)}}\right] = }$$ Tu mešana vsota odpade od pogojem, da imamo **neodvisne meritve**. Tako se račun nadaljuje:

$$= \frac{1}{N^2}N\sigma^2 = \frac{\sigma^2}{N}$$

Kar smo izračunali je pravzaprav tole:

$$\frac{dP}{dz_i} = \mathcal{N}(x, \sigma) \quad \Rightarrow \quad \frac{d\mathcal{P}}{d\bar{z}} = \mathcal{N}\left(x, \frac{\sigma}{\sqrt{\mathcal{N}}}\right)$$

Prevedeno iz matematičnega zapisa v jezik to pomeni, da smo dobili povprečje **porazdeljeno okoli istega x z ožjo Gaussovko.**

Videli smo, da so dodatne informacije izboljšale naše znanje o negotovosti spremenljivke. Sedaj pa si poglejmo to situacijo. Recimo, da smo merili v dveh "izmenah" in na koncu vsake izmene izračunali povprečje. Torej, da imamo:

$$N\,\mathrm{meritev} \;\Rightarrow\; \bar{z}_1 = \frac{1}{N}\sum_1^N z_i \quad \sigma_1 = \frac{\sigma}{\sqrt{N}}$$

$$M\,\mathrm{meritev} \;\Rightarrow\; \bar{z}_2 = \frac{1}{M}\sum_1^M z_i \quad \sigma_2 = \frac{\sigma}{\sqrt{M}}$$

Kako moramo utežiti te dve povprečji, da ju bomo **optimalno** združili? Poglejmo, kako bi zgledalo, če bi nekdo naredil kar

$N + M$ meritev. Tako bi imeli:

$$N + M \text{ meritev} \Rightarrow \bar{z}_3 = \frac{1}{N+M} \sum_{1}^{N+M} z_i \quad \sigma_3 = \frac{\sigma}{\sqrt{N+M}}$$

Vsoto v povprečju lahko ločimo na dva dela, tako da dobimo sledeče:

$$\bar{z}_3 = \frac{1}{N+M} \left[\sum_{1}^{N} z_i + \sum_{N+1}^{N+M} z_i \right] =$$

$$= \left(\frac{N}{N+M} \right) \bar{z}_1 + \left(\frac{M}{N+M} \right) \bar{z}_2$$

Opazimo, da smo dobili uteženo povprečje z dvema utzežema. Dajmo jih izraziti našimi variancami:

$$N = \frac{\sigma^2}{\sigma_1^2} \qquad M = \frac{\sigma^2}{\sigma_2^2}$$

$$N + M = \frac{\sigma^2}{\sigma_3^2} = \frac{\sigma^2}{\sigma_1^2} + \frac{\sigma^2}{\sigma_2^2}$$

$$\Rightarrow \frac{1}{\sigma_3^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

Torej se naša varianca ostri. Izraženo z variancami je potem optimalno združevanje:

$$\bar{z}_3 = \frac{\sigma_3^2}{\sigma_1^2} \bar{z}_1 + \frac{\sigma_3^2}{\sigma_2^2} \bar{z}_2 =$$

$$= \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \right) \bar{z}_1 + \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \right) \bar{z}_2$$

To lahko tudi zapišemo rekurzivno. Zadnji razliki pravimo “inovacija”.

$$z_3 = z_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} (z_2 - z_1)$$

Kvadratna forma

Do tega istega spoznanja lahko pridemo tudi preko kvadratne forme $2\dot{J}(x)$. Imejmo dve meritvi:

$$z_1 = \mathcal{N}(x, \sigma_1) \qquad z_2 = \mathcal{N}(x, \sigma_2)$$

To lahko pretvorimo v **Standardno Normalno Gaussovo porazdelitev** tako, da od spremenljivke odštejemo povprečje in delimo z varianco. Torej, da dobimo takole:

$$\frac{(z_1 - x)}{\sigma_1} = \mathcal{N}(0, 1) \qquad \frac{(z_2 - x)}{\sigma_2} = \mathcal{N}(0, 1)$$

Kvadratno formo zapišemo potem kot:

$$2J(x) = \frac{(z_1 - x)^2}{\sigma_1^2} + \frac{(z_2 - x)^2}{\sigma_2^2}$$

Ker hočemo optimalno združiti meritvi, želimo minimizirati seštevek. Zahtevamo:

$$\frac{d}{dx}(2J(x)) = 0$$

$$-\frac{2(z_1 - x)}{\sigma_1^2} - \frac{2(z_2 - x)}{\sigma_2^2} = 0$$

$$x\left[\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right] = \frac{z_1}{\sigma_1^2} + \frac{z_2}{\sigma_2^2}$$

Torej je optimalno združevanje:

$$x = z_3 = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-1}\left[\frac{z_1}{\sigma_1^2} + \frac{z_2}{\sigma_2^2}\right]$$

Disperzija optimalno združene ocene

Kaj če bi preverili ali je naše optimalno združevanje res optimalno?

$$z_1 = x + r_1 \quad \langle r_1 \rangle = 0 \quad \langle r_1^2 \rangle = \sigma_1^2$$

$$z_2 = x + r_2 \quad \langle r_2 \rangle = 0 \quad \langle r_2^2 \rangle = \sigma_2^2$$

Sedaj dajmo sestaviti z_3 kot linearno kombinacijo:

$$z_3 = \alpha z_1 + \beta z_1 = x + r$$

$$z_3 = \alpha(x + r_1) + \beta(x + r_2) = x + r$$

$$= (\alpha + \beta)x + \alpha r_1 + \beta r_2 = x + r$$

Od tod dobimo pogoja, da $\alpha + \beta = 1$ in:

$$r = \alpha r_1 + (1 - \alpha) r_2$$

Za skupen merilni šum seveda velja:

$$\langle r \rangle = 0$$

$$\langle r^2 \rangle = \alpha^2 \langle r_1^2 \rangle + (1 - \alpha)^2 \langle r_2^2 \rangle + 2\alpha(1 - \alpha)\langle r_1 r_2 \rangle$$

Od predpostavki, da merilna šuma meritev nista korelirana, zadnji člen odpade in ostane:

$$\langle r^2 \rangle = \alpha^2 \sigma_1^2 + (1 - \alpha)^2 \sigma_2^2$$

To sedaj minimiziramo po α:

$$\frac{d}{d\alpha}\langle r^2 \rangle = 0$$

$$2\alpha\sigma_1^2 - 2(1 - \alpha)\sigma_2^2 = 0$$

$$\alpha(\sigma_1^2 + \sigma_2^2) = \sigma_2^2$$

$$\Rightarrow \alpha = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \qquad \beta = 1 - \alpha = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

Dobili smo iste uteži kot prej, torej **ja** naše optimalno združevanje je res optimalno.

Korelacija med izmerki/ocenami

Imejmo dva seta meritev x in y. Kot vedno do zdaj velja:

$$\bar{r}_x = \bar{r}_y = 0$$

$$\bar{r_x^2} = \sigma_x^2 \neq 0 \qquad \bar{r_y^2} = \sigma_y^2 \neq 0$$

Denimo sedaj, da **obstaja korelacija** (povezava) med šumi x in y meritev.

$$\overline{r_x r_y} \neq 0$$

Definirajmo **kovarianco** kot:

$$\sigma_{xy} = \overline{(x-\bar{x})(y-\bar{y})} = \rho\sigma_x\sigma_y$$

kjer je $|\rho| \leq 1$ **korelacijski koeficient**. Negativni ρ bi tu pomenil antikorelacijo. Ta isti postopek lahko izvedemo tudi drugače (tako da vstavimo povprečje). Recimo, da imamo N meritev, potem lahko naredimo takole:

$$\sigma_{xy} = \frac{1}{N}\sum_i (x-\bar{x})(y-\bar{y}) = \frac{1}{N}\sum_i (xy - \bar{x}y - x\bar{y} + \bar{x}\bar{y}) =$$

$$= \frac{1}{N}\sum xy - \frac{1}{N}\bar{x}\sum y - \frac{1}{N}\bar{y}\sum x + \frac{1}{N}\bar{x}\bar{y}\sum 1 =$$

$$= \overline{xy} - \bar{x}\bar{y} - \bar{x}\bar{y} + \bar{x}\bar{y}$$

$$\Rightarrow \sigma_{xy} = \overline{xy} - \bar{x}\cdot\bar{y} \qquad \rho = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$$

Združevanje koreliranih meritev/ocen

Imejmo dva seta meritev kjer je merilni šum, tu označen z w, koreliran. Torej, da imamo nekaj v slogu:

$$z_1 = x + w_1 \qquad z_2 = x + w_2$$

$$\langle w_1^2\rangle = \sigma_1^2 \quad \langle w_2^2\rangle = \sigma_2^2 \quad \langle w_1 w_2\rangle \neq 0$$

Dajmo torej šum ene meritve zapisati kot linearno kombinacijo šuma druge meritve in neodvisnega dela (torej pravzaprav šum razstavimo):

$$w_1 = \alpha w_2 + w \Leftrightarrow (z_1 - x) = \alpha(z_2 - x) + w$$

$$\langle w\rangle = 0 \quad \langle w^2\rangle = \sigma_w^2 \quad \langle w w_2\rangle = 0$$

Razstavljena dela sta nekorelirana. Dajmo to sedaj izraziti s kovarianco:

$$\rho\sigma_1\sigma_2 = \langle w_1 w_2\rangle = \langle(\alpha w_2 + w)w_2\rangle = \alpha\langle w_2^2\rangle + \langle w w_2\rangle = \alpha\sigma_2^2$$

$$\Rightarrow \alpha = \rho\frac{\sigma_1}{\sigma_2}$$

Izrazimo še varianco neodvisnega šuma iz disperzije prvega šuma:

$$\sigma_1^2 = \langle w_1^2\rangle = \langle(\alpha w_2 + w)^2\rangle = \alpha^2\langle w_2^2\rangle + \langle w^2\rangle + \alpha\langle w w_2\rangle$$

Ne pozabiti, tu pa zadnji člen odpade, saj smo namensko ločili šum na odvisen in neodvisen del. Dobimo slednje:

$$\sigma_1^2 = \alpha^2\sigma_2^2 + \sigma_w^2 = \rho^2\frac{\sigma_1^2}{\sigma_2^2}\sigma_2^2 + \sigma_w^2$$

$$\Rightarrow \sigma_w^2 = \sigma_1^2(1-\rho^2)$$

Sestavimo sedaj kvadratno formo $2J(x)$:

$$2J(x) = \left(\frac{w_2}{\sigma_2}\right)^2 + \left(\frac{w}{\sigma_w}\right)^2 = \frac{w_2^2}{\sigma_2^2} + \frac{(w_1 - \alpha w_2)^2}{\sigma_w^2} =$$

$$= \frac{w_2^2}{\sigma_2^2} + \frac{w_1^2 - 2\alpha w_1 w_2 + \alpha^2 w_2^2}{\sigma_w^2} =$$

Sedaj lahko vstavimo, kar smo prej izrazili, da dobimo:

$$\frac{w_2}{\sigma_2^2} + \frac{w_2^2 \rho^2 \frac{\sigma_1^2}{\sigma_2^2}}{\sigma_1^2(1-\rho^2)} + \frac{w_1^2}{\sigma_1^2(1-\rho^2)} - \frac{2\rho \frac{\sigma_1}{\sigma_2} w_1 w_2}{\sigma_1^2(1-\rho^2)} =$$

$$= \frac{w^2}{\sigma_2^2}\left(1 + \frac{\rho^2}{1-\rho^2}\right) + \frac{w_1^2}{\sigma_1^2(1-\rho^2)} - \frac{2\rho w_1 w_2}{\sigma_1 \sigma_2 (1-\rho^2)}$$

Tako dobimo kvadratno formo:

$$2J(x) = \frac{1}{1-\rho^2}\left(\frac{(z_2 - x)^2}{\sigma_2^2} + \frac{(z_1 - x)^2}{\sigma_1^2} - \frac{2\rho(z_1 - x)(z_2 - x)}{\sigma_1 \sigma_2}\right)$$

Ker želimo da je združevanje optimalno, spet potegnemo staro foro:

$$\frac{d}{dx}(2J(x)) = 0$$

$$-\frac{2(z_2 - x)}{\sigma_2^2} - \frac{2(z_1 - x)}{\sigma_1^2} + \frac{2\rho}{\sigma_1 \sigma_2}(2x - (z_1 + z_2)) = 0$$

$$\frac{z_2}{\sigma_2^2} + \frac{z_1}{\sigma_1^2} - \frac{\rho(z_1 + z_2)}{\sigma_1 \sigma_2} = x\left[\frac{1}{\sigma_2^2} + \frac{1}{\sigma_1^2} - \frac{2\rho}{\sigma_1 \sigma_2}\right]$$

Da dobimo optimalen $\hat{z} = x$ je predpis potem:

$$\hat{z} = \left[\frac{1}{\sigma_2^2} + \frac{1}{\sigma_1^2} - \frac{2\rho}{\sigma_1 \sigma_2}\right]^{-1}\left(\frac{z_2}{\sigma_2^2} + \frac{z_1}{\sigma_1^2} - \frac{\rho(z_1 + z_2)}{\sigma_1 \sigma_2}\right)$$

Mejni primeri

Za konec si poglejmo še par mejnih primerov.

Ni korelacije

Če med šumi ni korelaciji, torej se pravi, da je $\rho = 0$, dobimo formulo, ki smo jo izpeljali v prejšnjem poglavju.

Popolna korelacija

V primeru popolne korelacije, torej se pravi, da je $\rho = 1$, dobimo:

$$x = \left(\frac{\sigma_2^2 + \sigma_1^2 - 2\sigma_1\sigma_2}{\sigma_1^2 \sigma_2^2}\right)^{-1}\left[\frac{z_1\sigma_2^2 + z_2\sigma_1^2 - (z_1 + z_2)\sigma_1\sigma_2}{\sigma_1^2\sigma_2^2}\right] =$$

$$= (\sigma_2 - \sigma_1)^{-2}\left[z_1\sigma_2(\sigma_2 - \sigma_1) - z_2\sigma_1(\sigma_2 - \sigma_1)\right] =$$

$$= (\sigma_2 - \sigma_1)^{-2}(\sigma_2 - \sigma_1)^2 z_2 = z_2$$

$$\Rightarrow \hat{\sigma^2} = \sigma_2^2 = \sigma_1^2$$

Enaka disperzija

V tem primeru je lahko ρ poljuben, velja pa $\sigma = \sigma_1 = \sigma_2$. Tako dobimo:

$$x = (2-2\rho)^{-1}\left[(z_1+z_2) - \rho(z_1+z_2)\right] =$$

$$= \frac{1-\rho}{2(1-\rho)}(z_1+z_2) = \frac{z_1+z_2}{2}$$

To pa seveda poznamo, ker je to le običajno povprečevanje.