Adaptivna obdelava signalov v telekomunikacijah

4. del

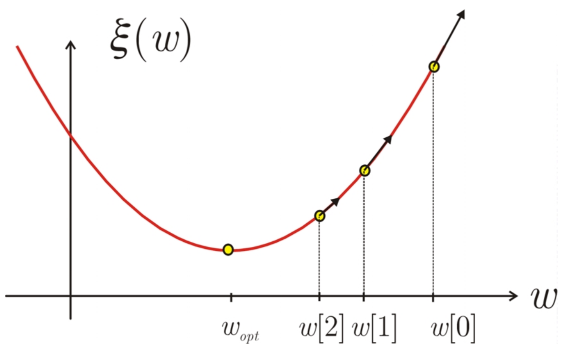
Pregled tem:

- Gradientni algoritem popravljanja uteži
  - potek popravljenja uteži
  - krivulja učenja
- Analiza gradientnega algoritma
  - presežek napake zaradi perturbacije uteži
  - napaka pri merjenju gradienta
  - presežek napake zaradi netočnosti gradienta

Gradientni algoritem (ponovitev)

- Utež popravljamo v nasprotni smeri gradienta kriterijske funkcije:

$$\boxed{w[m+1] = w[m] - \mu\nabla[m]}$$

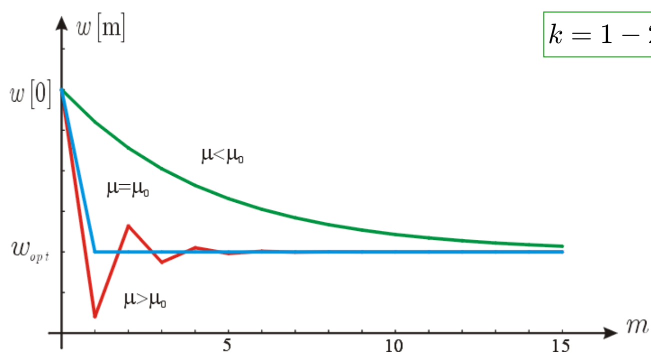


Potek popravljanja uteži

Kako izberemo vrednost adaptacijske konstante μ ?

$$\boxed{w[m] = w_{opt} + k^m(w[0] - w_{opt})}$$

$$\boxed{k = 1 - 2\mu R}$$

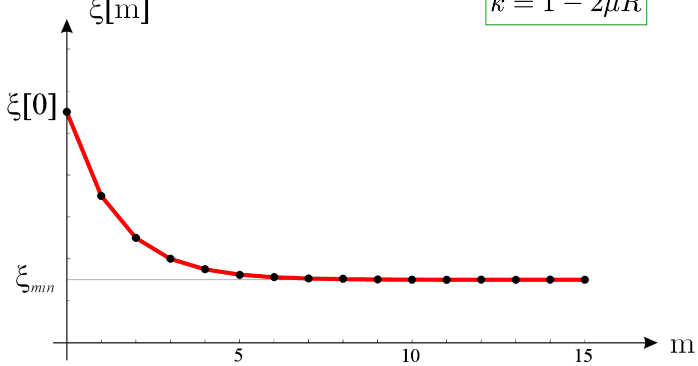


Krivulja učenja

- Potek zaporedja srednje kvadratične napake $\xi[m]$ imenujemo krivulja učenja

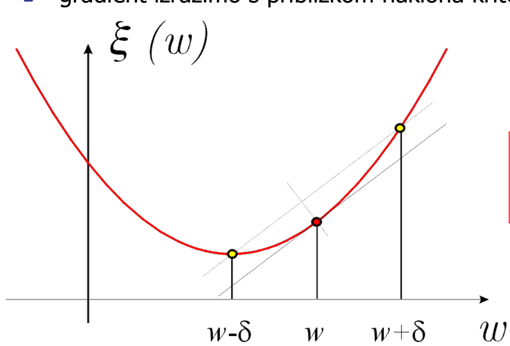
$$\boxed{\xi[m] = \xi_{min} + R(w[0] - w_{opt})^2 k^{2m}}$$

$$\boxed{k = 1 - 2\mu R}$$



Približek gradienta

- gradient izrazimo s približkom naklona kriterijske funkcije:

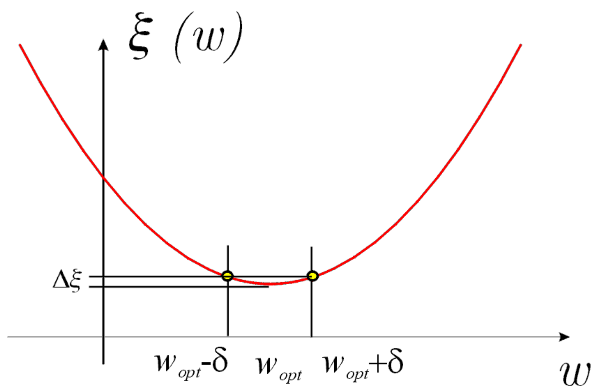


$$\boxed{\frac{d\xi}{dw} \approx \frac{\xi(w+\delta) - \xi(w-\delta)}{2\delta}}$$

- Gradient v točki w izračunamo na osnovi merjenja srednje kvadratične napake v dveh točkah z enakimo odmikom od w
- Ker natančne vrednosti kriterijske ne moremo izračunati na osnovi končnega števila meritev, je tudi gradient nenatančen !

Presežek napake zaradi perturbacije uteži

- Zaradi namernega spreminjanja uteži ($w+\delta$, $w-\delta$) se poveča srednja kvadratična vrednost napake tudi v okolici optimalne uteži ! Minimalna vrednost kriterijske funkcije je večja za $\Delta\xi$
- Relativno povečanje srednje kvadratične napake zaradi spreminjanja (perturbacije) uteži označimo s faktorjem neprilagojenosti M_p:

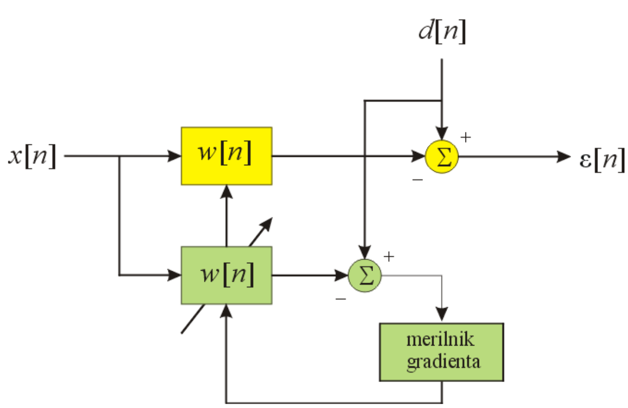


$$M_p = \frac{\Delta\xi}{\xi_{min}}$$

$$\boxed{M_p = \frac{\delta^2 R}{\xi_{min}}}$$

Postopek določitve približka gradienta:

- spreminjamo uteži na delujočem sistemu, posledica je povečanje kvadrata signala napake
- če se želimo izogniti povečanju napake moramo za meritev gradienta uporabiti vzporedni sistem:

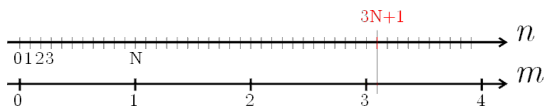


Napaka pri merjenju kriterijske funkcije

- Natančne vrednosti kriterijske ne moremo izračunati na osnovi končnega števila meritev !

$$\boxed{\xi = \overline{\varepsilon^2[n]} = \lim_{N\to\infty} \frac{1}{2N} \sum_{n=-N}^{N-1} \varepsilon^2[n]}$$

- Izračun vrednosti kriterijske funkcije na osnovi *N* meritev:



$$\boxed{\xi[m] = \frac{1}{N} \sum_{n=mN}^{N(m+1)-1} \varepsilon^2[n]}$$

- Napaka v izračunu kriterijske funkcije: $\boxed{\Delta\xi[m] = \xi[m] - \xi}$

Varianca kriterijske funkcije

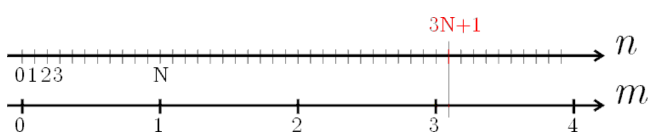
- Merilo napake pri merjenju kriterijske funkcije je varianca:

$$\mathrm{var}(\xi[m]) = \overline{(\Delta\xi[m])^2} = \overline{(\xi[m] - \xi)^2}$$

- Varianca kriterijske funkcije se zmanjša s povečanjem števila meritev *N* :

$$\mathrm{var}(\xi[m]) = \frac{\mathrm{var}(\varepsilon^2[n])}{N}$$

- S povečanjem števila meritev *N* se podaljšuje čas adaptacije !



Varianca kvadrata napake

- varianca trenutne vrednosti kvadrata napake:

$$\mathrm{var}(\varepsilon^2[n]) = \overline{(\varepsilon^2[n]-\xi)^2}$$

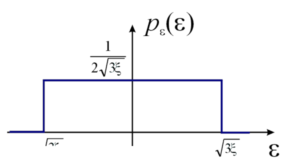
$$\boxed{\mathrm{var}(\varepsilon^2[n]) = \overline{\varepsilon^4[n]} - \xi^2}$$

- Vrednost četrtega momenta signala napake je odvisna od verjetnostne porazdelitve *p(ε)* :

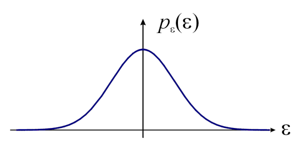
$$\overline{\varepsilon^4[n]} = \int_{-\infty}^{\infty} \varepsilon^4 p_\varepsilon(\varepsilon)\, d\varepsilon$$

Napaka pri merjenju kriterijske funkcije

- analiziramo varianco napake kriterijske funkcije za enakomerno in za normalno verjetnostno porazdelitev:



$$\overline{\varepsilon^4} = \frac{9\xi^2}{5}$$



$$\boxed{\overline{\varepsilon^4} = 3\xi^2}$$

- Četrti moment naključne spremenljivke je ravno največji pri Gaussovi porazdelitvi. Najbolj pesimistična ocena variance kvadrata napake je zato:

$$\mathrm{var}(\varepsilon^2[n]) = 3\xi^2 - \xi^2 = 2\xi^2$$

- Ocena variance kriterijske funkcije:

$$\boxed{\mathrm{var}(\xi[m]) = \frac{\mathrm{var}(\varepsilon^2[n])}{N} = \frac{2\xi^2}{N}}$$

Napaka gradienta

- Ker na osnovi končnega števila meritev ne moremo izračunati natančne vrednosti kriterijske funkcije, je tudi gradient nenatančen !

$$\hat{\nabla}[m] = \frac{\xi[m]_{(w+\delta)} - \xi[m]_{(w-\delta)}}{2\delta}$$

- napaka gradienta:

$$\eta[m] = \nabla[m] - \hat{\nabla}[m]$$

- varianca gradienta:

$$\boxed{\overline{\eta^2[m]} = \mathrm{var}(\hat{\nabla}[m]) = \frac{\xi^2}{N\delta^2}}$$

Varianca uteži od optimalne vrednosti

- Optimalne uteži praktično ne dosežemo zaradi nenatančnosti merjenja kriterijske funkcije in s tem tudi gradienta, kar povzroči dodatni šum v adaptacijskem procesu. Utež se stalno spreminja tudi v okolici optimalne vrednosti.
- Utež se spreminja v obe smeri okrog optimalne vrednosti. Merilo napake nastavitve uteži je varianca:

$$\boxed{\mathrm{var}(w) = \frac{\mu^2 \xi^2_{min}}{4N\delta^2 R(1-\mu R)}}$$

- Varianca uteži se zmanjša v primeru:
  - če povečamo N
  - če izberemo manjšo adaptacijsko konstanto μ
- Varianca uteži je obratnosorazmerna hitrosti adaptacije !

Presežek napake zaradi netočnosti gradienta

- Odstopanje kriterijske funkcije od minimalne vrednosti določa enačba:

$$\xi - \xi_{min} = v^2 R$$

- Zaradi variacije uteži nastopi v povprečju povečanje kvadrata napake:

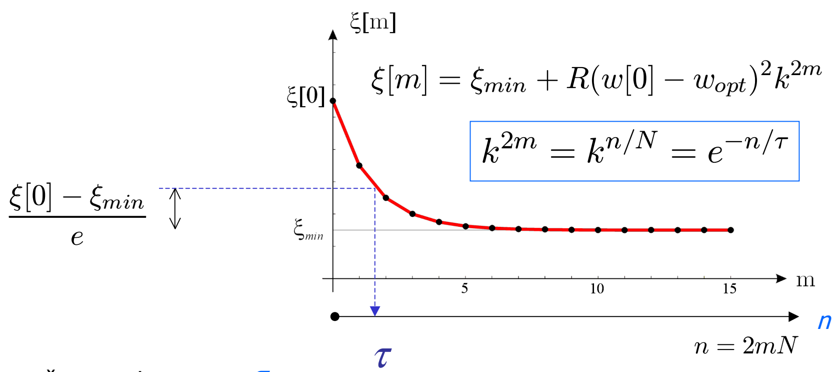
$$\overline{\xi - \xi_{min}} = \frac{\mu^2 \xi_{min}^2}{4N\delta^2(1-\mu R)}$$

  - Relativni presežek napake v okolici optimalne uteži je enak:

$$\boxed{M_g = \frac{\overline{\xi - \xi_{min}}}{\xi_{min}} = \frac{\mu \xi_{min}}{4N\delta^2(1-\mu R)}}$$

Časovna konstanta τ

- Zanima nas število vzorcev n, pri katerem pade presežek srednje kvadratične napake za faktor e. Odgovor poiščemo iz krivulje učenja:



- časovna konstanta τ:

$$\boxed{\tau = -\frac{N}{\ln k} \approx -\frac{N}{k-1} = \frac{N}{\mu R}}$$

Povečanje ξ zaradi perturbacije in napake gradienta

1. zaradi namerne perturbacije uteži pri meritvi gradienta $M_p = \frac{\delta^2 R}{\xi_{min}}$
2. zaradi napake pri meritvi gradienta $M_g = \frac{\xi_{min}}{4\delta^2 \tau R}$

$$\boxed{\xi_0 = (1 + M_p + M_g)\ \xi_{min}}$$

- neprilagojenosti sta obratnosorazmerni: $\boxed{M_g = \frac{1}{4\tau M_p}}$
- želimo zmanjšati oba vpliva hkrati:

$$\boxed{\xi_{0min} = (1 + \frac{1}{\sqrt{\tau}})\ \xi_{min}}$$

pogoj: $M_g = M_p = \frac{1}{2\sqrt{\tau}}$

Povezava minimalne napake s hitrostjo adaptacije

- srednja kvadratična napaka se v ravnotežnem stanju poveča za najmanj:

$$\boxed{\xi_{0min} = (1 + \frac{1}{\sqrt{\tau}})\ \xi_{min}}$$

Ne moremo imeti hkrati majhnega povečanja napake in hitre adaptacije uteži:

- hitra adaptacija pomeni manjhen τ in s tem večji ξ_{0min}
- počasna adaptacija pomeni velik τ in s tem manjši ξ_{0min}

Rešitev problema je v spreminjanju adaptacijske konstante :

- na začetku izberemo večjo vrednost adaptacijske konstante μ
- ko se približujemo optimalni vrednosti uteži, zmanjšujemo μ

$$\boxed{\tau = -\frac{N}{\ln k} \approx -\frac{N}{k-1} = \frac{N}{\mu R}}$$