

# Quantizzatore ad autoapprendimento

Enrico Odetti

Andrea Odetti\*

Dicembre 1967

Come noto un quantizzatore è un sistema dinamico senza memoria che trasforma un segnale analogico variabile d'ingresso in un segnale d'uscita discreto o digitale. Il diagramma più comune di un quantizzatore si presenta con gradini orizzontali (Fig. 1). Il numero dei gradini  $N$  può essere illimitato (quantizzatore senza saturazione).

Se  $N$  è limitato, come nel caso più comune, si definisce allora quantizzatore con saturazione. Nel caso più comune si definiscono solo l'altezza e la larghezza dei gradini (quantizzatore uniforme). Molto più interessante è quando l'altezza e la larghezza dei gradini sono una funzione del segnale d'ingresso o di una sua caratteristica statistica (quantizzatore non uniforme).

In questo lavoro studiamo un quantizzatore non uniforme con saturazione a cui applichiamo la funzione di autoapprendimento per ottimizzarne il comportamento.

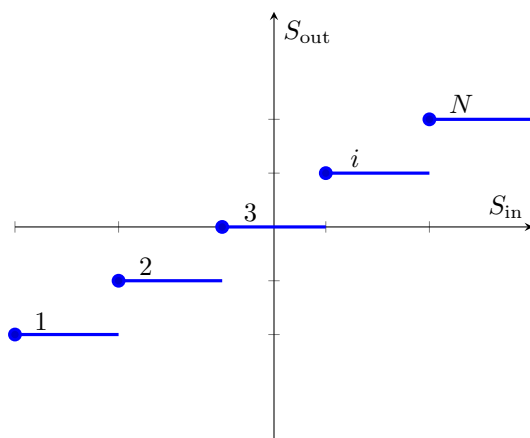
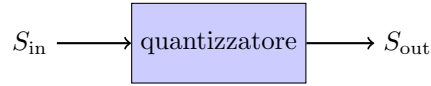


Figura 1: Diagramma di un quantizzatore uniforme.

Per  $N$  gradini il processo di quantizzazione è completamente definito se ai punti  $x_i$  e  $x_{i+1}$  di ciascun intervallo corrisponde un unico livello o gradino  $y_i$  (Fig. 2).

---

\*per la versione L<sup>A</sup>T<sub>E</sub>X Febbraio 2016



Per la migliore determinazione delle grandezze  $x_i$  e  $y_i$  introduciamo la misura di distribuzione  $D$ , per tutto il quantizzatore, come speranza matematica di una opportuna funzione  $f$  differenza tra il segnale d'ingresso e quello di uscita.

$$D = M\{f(S_{\text{in}} - S_{\text{out}})\} \quad (1)$$

Se la distribuzione di densità della probabilità di errore ( $S_{\text{in}} - S_{\text{out}}$ ) è uguale alla densità di distribuzione  $p(x)$  del segnale d'ingresso, possiamo scrivere

$$D = \sum_{i=1}^N \int_{x_i}^{x_{i+1}} f(x - y_i) p(x) dx \quad (2)$$

dove  $x_1 = -\infty, x_{N+1} = \infty$  e  $x$  è il segnale di entrata.

I migliori valori  $x_i$  e  $y_i$  devono minimizzare  $D$

$$\min_{x_i, y_i} D \rightarrow x_i^*, y_i^* \quad (3)$$

Nel caso la densità di distribuzione sia conosciuta in anticipo, il problema può essere considerato risolto. Se invece per quanto possibile non è nota, diviene allora complicata e complessa da calcolare. Nel secondo caso quando la funzione distribuzione  $p(x)$  non è nota, possiamo utilizzare un procedimento di auto-sintonia (quantizzatore ad autoapprendimento).

Supponiamo che le grandezze  $x_i$  e  $y_i$  possano assumere qualunque valore iniziale e che i valori successivi siano

$$\begin{aligned} x_i[0] & \quad i = 2, \dots, N \\ y_i[0] & \quad i = 1, \dots, N \end{aligned}$$

Un quantizzatore con sintonizzatore si realizza con un dispositivo speciale, che chiamiamo *adattatore* (Fig. 3). L'adattatore, per passi successivi, calcola i valori  $x_i$  e  $y_i$  per cui si ottengono i valori ottimali  $x_i^*$  e  $y_i^*$ .

Supponiamo che le variazioni si effettuino in modo discreto nel tempo, in tal modo possiamo scrivere la corrispondenza tra ingresso e uscita come segue

$$\left. \begin{aligned} x[n] \\ x_i[n-1], y_i[n-1] \end{aligned} \right\} \rightarrow x_i[n], y_i[n] \quad (4)$$

In questo lavoro si stabilisce la corrispondenza ingresso-uscita dell'adattatore (4) formalmente con i principi dell'approssimazione stocastica. Come noto il metodo dell'approssimazione stocastica permette di trovare i valori estremi del funzionale (2) quando la conoscenza a priori della densità di distribuzione  $p(x)$  soddisfa qualche condizione che inizialmente non conosciamo. Per minimizzare

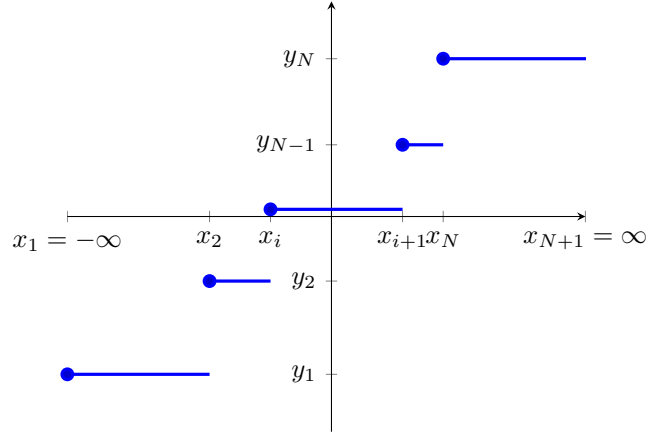


Figura 2: Diagramma di un quantizzatore non uniforme.

$D$  per  $N$  costante differenziamo per  $x_i$  e  $y_i$  e imponiamo che il risultato abbia valore nullo

$$\frac{\partial D}{\partial x_i} = f(x_i - y_{i-1})p(x_i) - f(x_i - y_i)p(x_i) = 0 \quad i = 2, \dots, N \quad (5)$$

$$\frac{\partial D}{\partial y_i} = - \int_{x_i}^{x_{i+1}} f(x - y_i)p(x) dx = 0 \quad i = 1, \dots, N \quad (6)$$

Dalla (5) otteniamo (quando  $p(x_i) \neq 0$ )

$$f(x_i - y_{i-1}) = f(x_i - y_i) \quad i = 2, \dots, N \quad (7)$$

È evidente che l'ultima soluzione non dipende da  $p(x)$  al contrario della soluzione (6). Considerando la soluzione (7), la condizione di minimo può essere scritta come segue

$$\begin{cases} f(x_i - y_{i-1}) = f(x_i - y_i) \\ \min_{y_i} D \end{cases} \rightarrow x_i^*, y_i^* \quad (8)$$

dove adesso il minimo è ricercato solo per  $y_i$  ( $i = 1, \dots, N$ ). La soluzione (7) si presenta come una condizione supplementare.

Supponendo che ci sia solo un sistema risolutivo (8) possiamo trovarlo con il metodo dell'approssimazione stocastica.

Indichiamo

$$\bar{c} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (9)$$

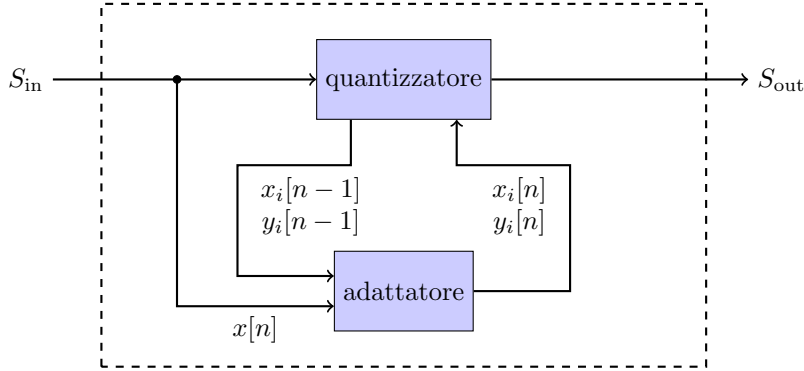


Figura 3: Struttura di un quantizzatore ad autoapprendimento.

Per cui

$$\min_{\bar{c}} D(\bar{c}) = \min_{\bar{c}} M_x \{f(\bar{c}, x)\} \rightarrow \bar{c}^* \quad (10)$$

Non scriviamo una nuova condizione (7), in rapporto a quelle supposte, poiché si realizzano sempre.  $\bar{c}^*$  deve risolvere la seguente equazione

$$\Delta D(\bar{c}) = M_x \{\nabla_{\bar{c}} f(\bar{c}, x)\} = 0 \quad (11)$$

Possiamo risolvere l'equazione precedente con il seguente algoritmo di approssimazione stocastica

$$\bar{c}[n] = \bar{c}[n-1] - \gamma[n] \nabla_{\bar{c}} f(x[n], \bar{c}[n-1]) \quad (12)$$

dove  $\gamma[n]$  è una successione di valori favorevoli, tali per cui

$$\sum_{n=1}^{\infty} \gamma[n] = \infty \quad \sum_{n=1}^{\infty} \gamma^2[n] < \infty \quad (13)$$

In tal modo affinché di fatto siano determinati  $x_i$  e  $y_i$  poniamo  $f(\bullet)$  nella forma di una funzione quadratica. Dopo aver calcolato  $\nabla_{\bar{c}} f$  otteniamo

$$\frac{\partial f}{\partial y_i} = -2(x - y_i) \varepsilon_{x_i, x_{i+1}}(x) \quad i = 1, \dots, N \quad (14)$$

dove

$$\varepsilon_{x_i, x_{i+1}}(x) = \begin{cases} 1 & x_i \leq x < x_{i+1} \\ 0 & \text{altrove} \end{cases}$$

Adesso possiamo scrivere dettagliatamente l'algoritmo (12) prendendo in considerazione la condizione (7)

$$\begin{aligned}
f(x_i - y_{i-1}) &= f(x_i - y_i) \\
(x_i - y_{i-1})^2 &= (x_i - y_i)^2 \\
x_i &= \frac{y_i + y_{i-1}}{2} \\
i &= 2, \dots, N
\end{aligned} \tag{15}$$

oppure

$$\begin{cases}
x_2[n] = \frac{y_1[n] + y_2[n]}{2} \\
\cdots \\
x_N[n] = \frac{y_{N-1}[n] + y_N[n]}{2} \\
\begin{bmatrix} y_1[n] \\ \vdots \\ y_N[n] \end{bmatrix} = \begin{bmatrix} y_1[n-1] \\ \vdots \\ y_N[n-1] \end{bmatrix} + 2\gamma[n] \begin{bmatrix} (x[n] - y_1[n-1])\varepsilon_{x_1, x_2[n-1]}(x[n]) \\ \vdots \\ (x[n] - y_N[n-1])\varepsilon_{x_N[n-1], x_{N+1}}(x[n]) \end{bmatrix}
\end{cases} \tag{12'}$$

dove sempre  $x_1 = -\infty$ ,  $x_{N+1} = \infty$ .

Come noto da  $\gamma[n]$  dipende la convergenza, come anche la velocità del processo di convergenza dell'autoapprendimento.

Per questo affinché di fatto sia definita una unica funzione  $\gamma[n]$ , impostiamo una condizione supplementare così che il processo iterativo converga nel modo migliore.

Il criterio di scelta della qualità dell'algoritmo si determina come segue: il funzionale  $D(\bar{c}) = M_x \{(S_{in} - S_{out})^2\}$  si imposta empiricamente

$$V_e^2[n] = \frac{1}{n} \sum_{m=1}^n \sum_{i=1}^N (x[m] - y_i[n])^2 \varepsilon_{x_i[n-1], x_{i+1}[n-1]}(x[m]) \tag{16}$$

otteniamo

$$\min_{\gamma[n]} V_e^2[n] \rightarrow \gamma^*[n] \tag{17}$$

cioè

$$\frac{dV_e^2[n]}{d\gamma[n]} = 0 \tag{18}$$

Sostituendo nella (16)  $y_i[n]$ , nell'algoritmo (12'), otteniamo

$$\begin{aligned}
V_e^2[n] &= \frac{1}{n} \sum_{m=1}^n \sum_{i=1}^N \left( x[m] - \left( y_i[n-1] + 2\gamma[n] (x[n] - y_i[n-1]) \varepsilon_{x_i[n-1], x_{i+1}[n-1]}(x[n]) \right) \right)^2 \\
&\quad \times \varepsilon_{x_i[n-1], x_{i+1}[n-1]}(x[m])
\end{aligned}$$

Di conseguenza l'equazione (18) diventa

$$\frac{dV_e^2[n]}{d\gamma[n]} = -\frac{2}{n} \sum_{m=1}^n \sum_{i=1}^N \left( x[m] - \left( y_i[n-1] + 2\gamma[n] (x[n] - y_i[n-1]) \right) \right) \times \\ (x[n] - y_i[n-1]) \varepsilon_{x_i[n-1], x_{i+1}[n-1]}(x[n]) \varepsilon_{x_i[n-1], x_{i+1}[n-1]}(x[m]) = 0 \quad (18')$$

Nella (18')  $\varepsilon(x[n])$  garantiscono l'esistenza di un solo valore sommatoria  $\sum_{i=1}^N$ . Con questo si completa l'equazione

$$\sum_{m=1}^n \left( x[m] - \left( y_{\underline{i}}[n-1] + 2\gamma[n] (x[n] - y_{\underline{i}}[n-1]) \right) \right) (x[n] - y_{\underline{i}}[n-1]) = 0 \quad (19)$$

oppure

$$\sum_{m=1}^n \left( x[m] - \left( y_{\underline{i}}[n-1] + 2\gamma[n] (x[n] - y_{\underline{i}}[n-1]) \right) \right) = 0$$

dove  $\underline{i}$  sono determinati dalla condizione

$$\varepsilon_{x_{\underline{i}}[n-1], x_{\underline{i}+1}[n-1]}(x[n]) \neq 0 \quad (20)$$

da cui otteniamo il valore ottimale per  $\gamma[n]$

$$\gamma^*[n] = \frac{\frac{1}{n^*} \sum_{m=1}^n x[m] - y_{\underline{i}}[n-1]}{2(x[n] - y_{\underline{i}}[n-1])} \quad (21)$$

dove la sommatoria dei valori  $x[m]$ , che appartengono all'intervallo  $x_{\underline{i}}[n-1], x_{\underline{i}+1}[n-1]$  per i quali, per ogni  $\varepsilon_{x_{\underline{i}}[n-1], x_{\underline{i}+1}[n-1]}(x[m]) \neq 0$  e  $n^*$  è il numero di questi valori.

Se introduciamo  $\gamma^*[n]$  nell'algoritmo, otteniamo

$$y_i = \frac{1}{n^*} \sum_{m=1}^n x[m] \quad (22)$$

Questa equazione mostra che per ogni passo i valori  $y_i^*$  tendono a eguagliare il valore medio di tutti gli  $x[m]$ , che appartengono all'intervallo  $x_i[n-1], x_{i+1}[n-1]$  al passo precedente.

Al termine di questo lavoro<sup>1</sup> desidero ringraziare caldamente il prof. Yakov Zalmanovitch Tsypkin<sup>2</sup> per il grande aiuto e appoggio datomi per la sua stesura.

<sup>1</sup>Questo lavoro costituisce la tesi di fine corso, giugno 1967, di uno stage presso l'Istituto di Energetica di Mosca MEI (<http://www.mpei.ru>), annesso al Politecnico.

<sup>2</sup>In Memoria di Yakov Zalmanovitch Tsypkin: A Life in Feedback Control, <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=664147>

## Riferimenti bibliografici

- [1] Цыпкин Я. З., Адаптация, обучение и самообучение в автоматических системах, Автоматика и телемеханика, 1, 1966.
- [2] Joel Max, Quantizing for minimum distortion, IRE Transactions on Information Theory, Vol. 6 marzo 1960, N. 1.
- [3] Кошелев В. Н., Квантование с минимальной энтропией, Проблемы передачи информации АН СССР, выпуск 14, 1963.
- [4] Солодов А. В., Теория информации и ее применение к задачам автоматического управления и контроля, Изд-во Наука, Москва, 1967.
- [5] Tou J., Optimum design of digital control systems, Academic Press, 1963.