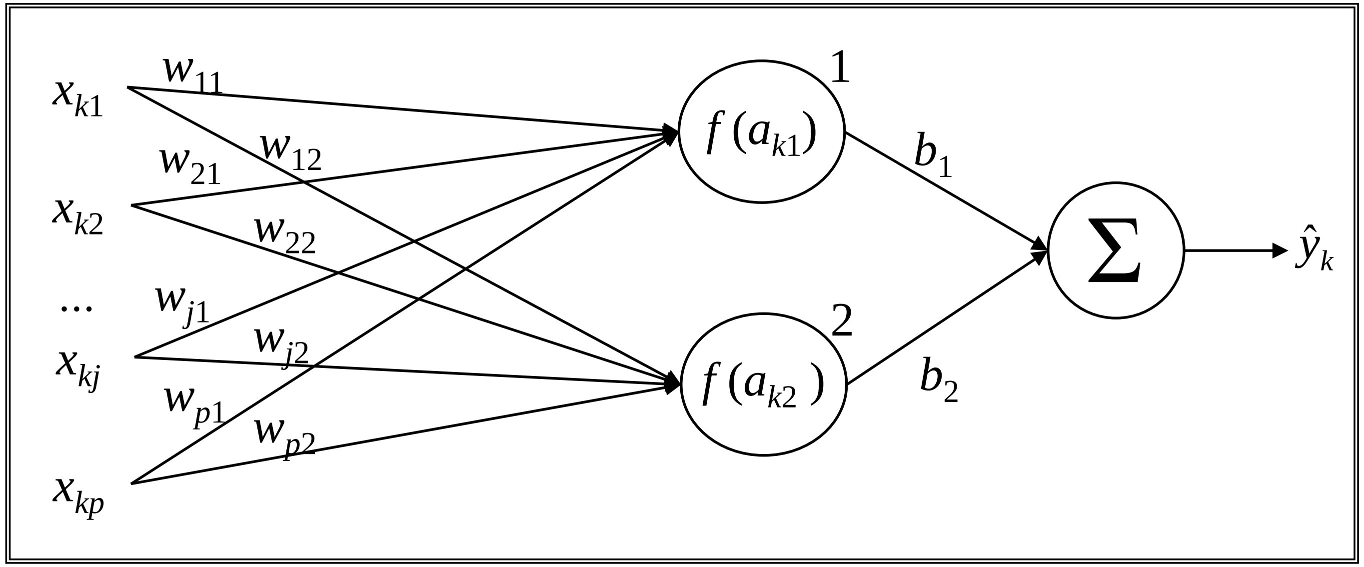
Μάθημα 9

Εκπαίδευση Νευρωνικών Δικτύων Εμπρόσθιας Τροφοδότησης (ΝΔΕΤ)

(ΣΥΝΕΧΕΙΑ)

Βασική Δομή Νευρωνικού Δικτύου Εμπρόσθιας Τροφοδότησης (ΝΔΕΤ) με Δύο Τεχνητούς Νευρώνες



Συνεπώς, η έξοδος του ΝΔΕΤ με 2 νευρώνες είναι:

$$\hat{y}_k = f(a_{k1})b_1 + f(a_{k2})b_2 = \sum_{i=1}^{2} f(a_{ki})b_i$$

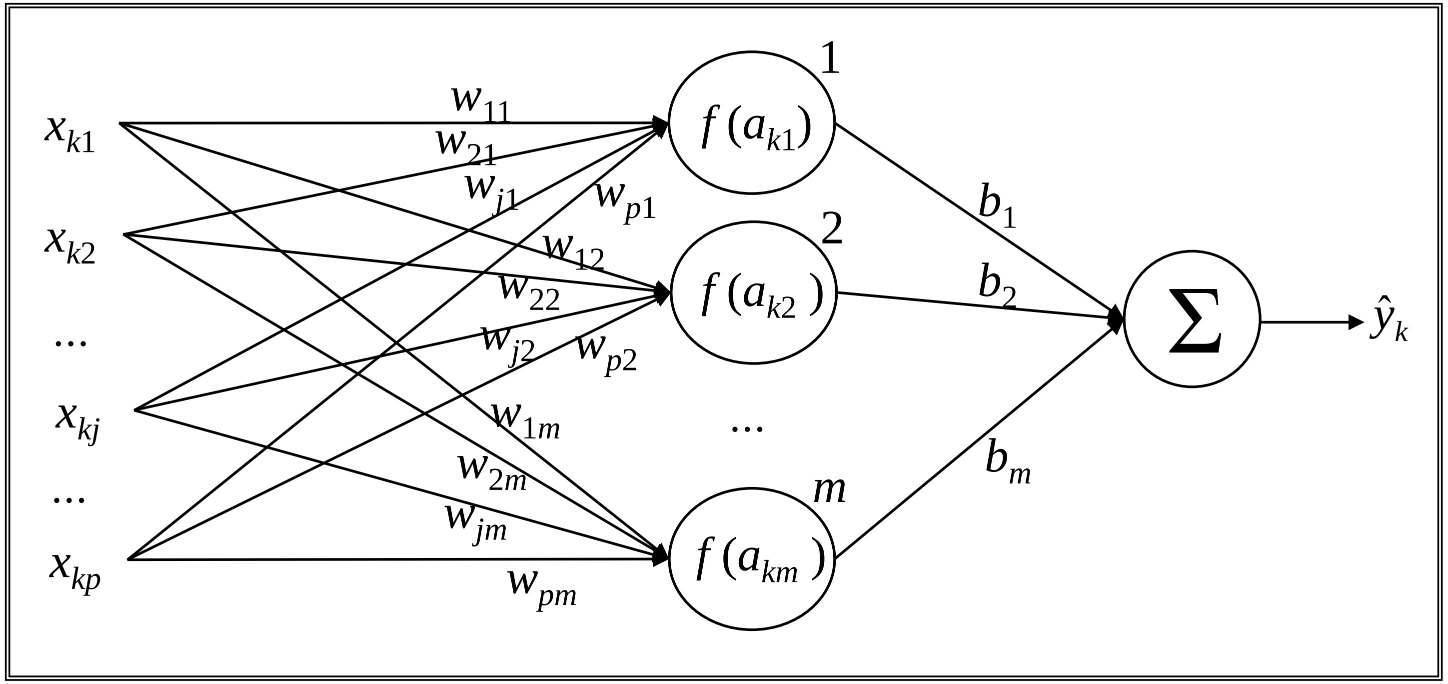
$$a_{ki} = x_{k1}w_{1i} + x_{k2}w_{2i} + \ldots + x_{kj}w_{ji} + \ldots + x_{kp}w_{pi} = \sum_{j=1}^{p} x_{kj}w_{ji}$$

$$\hat{y}_k = \sum_{i=1}^{2} f\left(\sum_{j=1}^{p} x_{kj}w_{ji}\right) b_i$$

$$f(a_{ki}) = \frac{1}{1+e^{-a_{ki}}} = \frac{1}{1+\exp(-a_{ki})}$$

$$\boxed{\hat{y}_k = \sum_{i=1}^{2} \frac{1}{1+\exp\left(-\sum_{j=1}^{p} x_{kj}w_{ji}\right)} b_i}$$

Βασική Δομή Νευρωνικού Δικτύου Εμπρόσθιας Τροφοδότησης (ΝΔΕΤ) με *m* Τεχνητούς Νευρώνες



Συνεπώς, η έξοδος του ΝΔΕΤ με *m* νευρώνες είναι:

$$\hat{y}_k = f(a_{k1})b_1 + f(a_{k2})b_2 + \ldots + f(a_{km})b_m = \sum_{i=1}^{m} f(a_{ki})b_i$$

$$a_{ki} = x_{k1}w_{1i} + x_{k2}w_{2i} + \ldots + x_{kj}w_{ji} + \ldots + x_{kp}w_{pi} = \sum_{j=1}^{p} x_{kj}w_{ji}$$

$$\hat{y}_k = \sum_{i=1}^{m} f\Big(\sum_{j=1}^{p} x_{kj}w_{ji}\Big)b_i$$

$$f(a_{ki}) = \frac{1}{1+e^{-a_{ki}}} = \frac{1}{1+\exp(-a_{ki})}$$

$$\boxed{\hat{y}_k = \sum_{i=1}^{m} \frac{1}{1+\exp\big(-\sum_{j=1}^{p} x_{kj}w_{ji}\big)} b_i}$$

Συμπερασματικά

Έξοδος του ΝΔΕΤ με 2 νευρώνες:

$$\hat{y}_k = \sum_{i=1}^{2} \frac{1}{1+\exp(-\sum_{j=1}^{p} x_{kj} w_{ji})} b_i$$

Έξοδος του ΝΔΕΤ με *m* νευρώνες:

$$\hat{y}_k = \sum_{i=1}^{m} \frac{1}{1+\exp(-\sum_{j=1}^{p} x_{kj} w_{ji})} b_i$$

Και στις δύο περιπτώσεις η συνάρτηση σφάλματος είναι:

$$E = \sum_{k=1}^{N} (y_k - \hat{y}_k)^2$$

Για λόγους στατιστικής ευχέρειας στο πρόγραμμα εμφανίζουμε συνήθως την ρίζα του μέσου τετραγωνικού σφάλματος

Root Mean Square Error (RMSE)

$$\boxed{RMSE = \sqrt{\frac{\sum_{k=1}^{N} (y_k - \hat{y}_k)^2}{N}}}$$

Εκπαίδευση Νευρωνικού Δικτύου Εμπρόσθιας Τροφοδότησης (ΝΔΕΤ) με *m* Τεχνητούς Νευρώνες
Μέθοδος Καθοδικής Κλίσης (Gradient Descent Method)

Αντικειμενική Συνάρτηση

$$E(w_{11}, w_{12}, ..., w_{pm}, b_1, b_2, ..., b_m) = \sum_{k=1}^{N} \left(y_k - \sum_{i=1}^{m} \frac{1}{1 + \exp(-\sum_{j=1}^{p} x_{kj} w_{ji})} b_i \right)^2$$

Μερικές Παράγωγοι

$$\frac{\partial E}{\partial w_{ji}} = (-2) \sum_{k=1}^{N} \left(y_k - \hat{y}_k \right) \frac{\exp(-\sum_{j=1}^{p} x_{kj} w_{ji})}{\left(1 + \exp(-\sum_{j=1}^{p} x_{kj} w_{ji}) \right)^2} x_{kj}\, b_i$$

$$\frac{\partial E}{\partial b_i} = (-2) \sum_{k=1}^{N} \left(y_k - \hat{y}_k \right) \frac{1}{1 + \exp(-\sum_{j=1}^{p} x_{kj} w_{ji})}$$

Κανόνες Εκμάθησης

$$w_{ji}^{new} = w_{ji}^{old} - \eta \frac{\partial E}{\partial w_{ji}^{old}} \Rightarrow \boxed{w_{ji} = w_{ji} - \eta \frac{\partial E}{\partial w_{ji}}}$$

$$b_i^{new} = b_i^{old} - \eta \frac{\partial E}{\partial b_i^{old}} \Rightarrow \boxed{b_i = b_i - \eta \frac{\partial E}{\partial b_i}}$$

Εκπαίδευση Νευρωνικού Δικτύου Εμπρόσθιας Τροφοδότησης (ΝΔΕΤ): Ανάλυση Δεδομένων

- Τα δεδομένα Εισόδου-Εξόδου χωρίζονται σε δύο σύνολα
  - Δεδομένα εκπαίδευσης: αποτελούν συνήθως το 60% των συνολικών δεδομένων
  - Δεδομένα αξιολόγησης: αποτελούν συνήθως το 40% των συνολικών δεδομένων
  - Τα δεδομένα εκπαίδευσης (και άρα τα δεδομένα αξιολόγησης συνήθως επιλέγονται τυχαία αλλά μία καλή στρατηγική είναι να επιλέξουμε ως δεδομένα εκπαίδευσης τα πρώτα 60% και ως δεδομένα αξιολόγησης τα τελευταία 40% των συνολικών δεδομένων
  - Το νευρωνικό δίκτυο δεν έχει δει καθόλου τα δεδομένα αξιολόγησης. Οπότε τα δεδομένα αξιολόγησης χρησιμοποιούνται για τον έλεγχο της αξιοπιστίας/εγκυρότητας του νευρωνικού δικτύου

$\boldsymbol{X}$

		x_1	x_2	…	x_j	…	x_p
Δεδομένα εισόδου εκπαίδευσης	1	x_{11}	x_{12}	…	x_{1j}	…	x_{1p}
	2	x_{21}	x_{22}	…	x_{2j}	…	x_{2p}
	…	…	…	…	…	…	…
	N_Train	$x_{N\_Train,1}$	$x_{N\_Train,2}$	…	$x_{N\_Train,j}$	…	$x_{N\_Train,p}$
Δεδομένα εισόδου αξιολόγησης	$N\_Train+1$	$x_{N\_Train+1,1}$	$x_{N\_Train+1,2}$	…	$x_{N\_Train+1,j}$	…	$x_{N\_Train+1,p}$
	…	…	…	…	…	…	…
	N	x_{N1}	x_{N2}	…	x_{Nj}	…	x_{Np}

$\boldsymbol{Y}$

		y
Δεδομένα εξόδου εκπαίδευσης	1	y_1
	2	y_2
	…	…
	N_Train	y_{N_Train}
Δεδομένα εξόδου αξιολόγησης	$N\_Train+1$	$y_{N\_Train+1}$
	…	…
	N	y_N

Εκπαίδευση Νευρωνικού Δικτύου Εμπρόσθιας Τροφοδότησης (ΝΔΕΤ): Ανάλυση Δεδομένων

Δεδομένα εισόδου-εξόδου εκπαίδευσης

X

	x_1	x_2	…	x_j	…	x_p
1	x_{11}	x_{12}	…	x_{1j}	…	x_{1p}
2	x_{21}	x_{22}	…	x_{2j}	…	x_{2p}
…	…	…	…	…	…	…
N_Train	$x_{N\_Train,1}$	$x_{N\_Train,2}$	…	$x_{N\_Train,j}$	…	$x_{N\_Train,p}$

Y

	y
1	y_1
2	y_2
…	…
N_Train	y_{N_Train}

N: Συνολικός αριθμός δεδομένων εισόδου-εξόδου

N_Train: Αριθμός δεδομένων εισόδου-εξόδου εκπαίδευσης

N_Test: Αριθμός δεδομένων εισόδου-εξόδου εκπαίδευσης

N=N_Train+N_Test

Δεδομένα εισόδου-εξόδου αξιολόγησης

X

	x_1	x_2	…	x_j	…	x_p
1	x_{11}	x_{12}	…	x_{1j}	…	x_{1p}
2	x_{21}	x_{22}	…	x_{2j}	…	x_{2p}
…	…	…	…	…	…	…
N_Test	$x_{N\_Train,1}$	$x_{N\_Train,2}$	…	$x_{N\_Train,j}$	…	$x_{N\_Train,p}$

Y

	y
1	y_1
2	y_2
…	…
N_Train	y_{N_Train}

Εκπαίδευση Νευρωνικού Δικτύου Εμπρόσθιας Τροφοδότησης (ΝΔΕΤ): Ανάλυση Δεδομένων

```
x_data=xlsread('x_Data.xlsx');
y_data=xlsread('y_Data.xlsx');
dim=size(x_data);
N=dim(1);
p=dim(2);
N_Train=round(0.6*N);
N_Test=N_Data-N;
for k=1:N_Train
    for j=1:p
       x(k,j)=x_data(k,j);
   end
   y(k)=y_data(k);
end;

for k=1:N_Test
    for j=1:p
       x_test(k,j)=x_data(N+k,j);
   end
   y_test(k)=y_data(N+k);
end;
```

Η συνάρτηση round(a) επιστρέφει τον πιο κοντινό ακέραιο στον αριθμό a

π.χ. round(3.4)=3,   round(6.55)=7

```
%---Δεδομένα Εισόδου-Εξόδου------
x=xlsread('x_Data.xlsx');
y=xlsread('y_Data.xlsx');
dim=size(x);
N=dim(1);
p=dim(2);
N_train=round(0.6*N);
N_test=N-N_train;
x_train=zeros(N_train,p);
y_train=zeros(N_train,1);
x_test=zeros(N_test,p);
y_test=zeros(N_test,1);

for k=1:N_train
  for j=1:p
    x_train(k,j)=x(k,j);
  end
  y_train(k)=y(k);
end;

for k=1:N_test
  for j=1:p
    x_test(k,j)=x(N_train+k,j);
  end
  y_test(k)=y(N_train+k);
end;

m=5;    %---Αριθμός νευρώνων
ni=0.00005;  %---Ρυθμός εκμάθησης
```

```
%----Τυχαία Αρχικοποίηση των Βαρών
for j=1:p
  for i=1:m
    w(j,i)=-0.001+rand*0.002;
  end
end

for i=1:m
  b(i)=-0.001+rand*0.002;
end
```

```
%---Υπολογισμός του y_est και της συνάρτησης σφάλματος----
y_est_train = Estimated_Output (N_train,p,m,x_train,w,b);
Error_train = Error_Function (N_train,y_train,y_est_train);
RMSE_train=sqrt(Error_train/N_train)

y_est_test = Estimated_Output(N_test,p,m,x_test,w,b);
Error_test = Error_Function (N_test,y_test,y_est_test);
RMSE_test=sqrt(Error_test/N_test)

%---Επαναληπτική Διαδικασία Μηχανικής Μάθησης-----
for it=1:200
  it=it
  [der_w,der_b] = Derivatives (N_train,p,m,x_train,y_train,y_est_train,w,b);
  for j=1:p
    for i=1:m
      w(j,i)=w(j,i)-ni*der_w(j,i);
    end
  end
  for i=1:m
    b(i)=b(i)-ni*der_b(i);
  end

  y_est_train = Estimated_Output (N_train,p,m,x_train,w,b);
  Error_train = Error_Function (N_train,y_train,y_est_train);
  RMSE_train=sqrt(Error_train/N_train)

  y_est_test = Estimated_Output(N_test,p,m,x_test,w,b);
  Error_test = Error_Function (N_test,y_test,y_est_test);
  RMSE_test=sqrt(Error_test/N_test)
end
```

Ασαφής Λογική

Ό όρος Ασαφής Λογικής αναφέρεται σε πράγματα, τα οποία εμπεριέχουν μία ασάφεια ή δεν είναι ρητά δηλωμένα. Στην πραγματική ζωή, εμείς οι άνθρωποι συνεχώς εκφραζόμαστε με αυτόν τον τρόπο
Π.χ.

- Ερ.: Τι καιρό κάνει; Απ.: Καλό καιρό
- Ερ. Κάνει κρύο; Απ.: Όχι πολύ.
- Πως είσαι; Απ. Μια χαρά.

Όλη αυτή η ανθρώπινη συλλογική περιγράφεται από την Ασαφή Λογική.

Ένας αλγόριθμος Ασαφούς Λογικής είναι μία υπολογιστική πλατφόρμα, η οποία

1. Εμπεριέχει την αβεβαιότητα που υπάρχει στα δεδομένα
2. Χρησιμοποιεί Κανόνες Αν…Τότε για να περιγράψει το πραγματικό σύστημα
3. Εξάγει συμπεράσματα με τον τρόπο που το κάνουν οι άνθρωποι.

Ιστορικά Στοιχεία

- Δεκαετία 1920: Οι πρώτες μελέτες για μία διαφορετική λογική από τον προτασιακό και κατηγορικό λογισμό

- 1965: Ο Lofti Zadeh (Καθηγητής στο Berkeley) δημοσιεύει το πρώτο άρθρο όπου εισάγει την έννοια του ασαφούς συνόλου

- Σήμερα: Είναι ένα από τα βασικά αντικείμενα της Υπολογιστικής Νοημοσύνης και η αιχμή του δόρατος στην σύγχρονη επιστήμη της Τεχνητής Νοημοσύνης

Βασικά Χαρακτηριστικά Στοιχεία

- Ευχρηστία και ευκολία στην υλοποίηση
- Βασίζεται στην μίμηση της ανθρώπινης λογικής και λογισμού
- Υψηλή απόδοση σε προβλήματα τα οποία εμπεριέχουν ασάφεια και αβεβαιότητα (π.χ. ιατρική διάγνωση)
- Επιτρέπει την δημιουργία μη-γραμμικών μοντέλων με μεγάλη πολυπλοκότητα
- Ένα ασαφές μοντέλο μπορεί να σχεδιαστεί όχι μόνο με την χρήση δεδομένων αλλά και με την βοήθεια ενός ειδικού

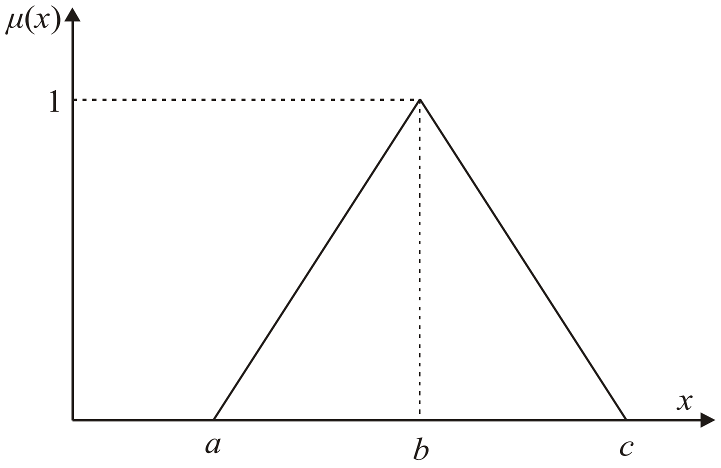
Πότε δεν Χρησιμοποιούμε Ασαφή Λογική

Η Ασαφής Λογική δεν θεραπεύει τα πάντα....Δεν χρησιμοποιούμε Ασαφή Λογική όταν

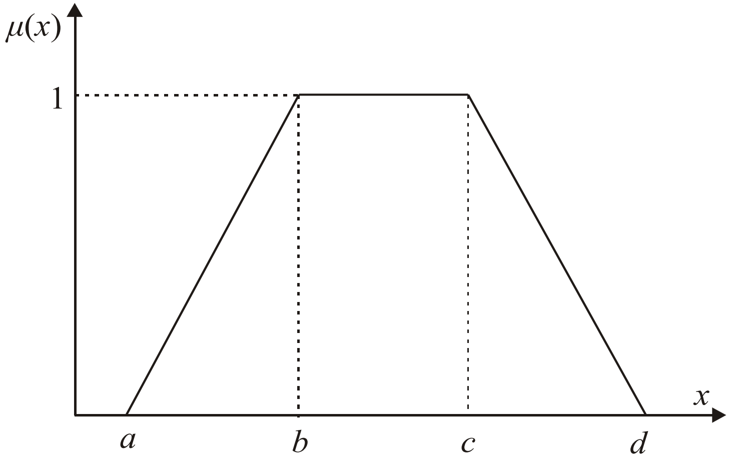
- Δεν υπάρχει λόγος να βρούμε με περιγραφικό τρόπο την σχέση μεταξύ δεδομένων εισόδου και δεδομένων εξόδου

- Μπορούμε να χρησιμοποιήσουμε κοινή λογική (δηλ. προτασιακή και κατηγορική λογική) με πολύ εύκολο τρόπο. Π.χ. σε μία ιατρική διάγνωση που απαιτείται ρητά να ειπωθεί κάτι…

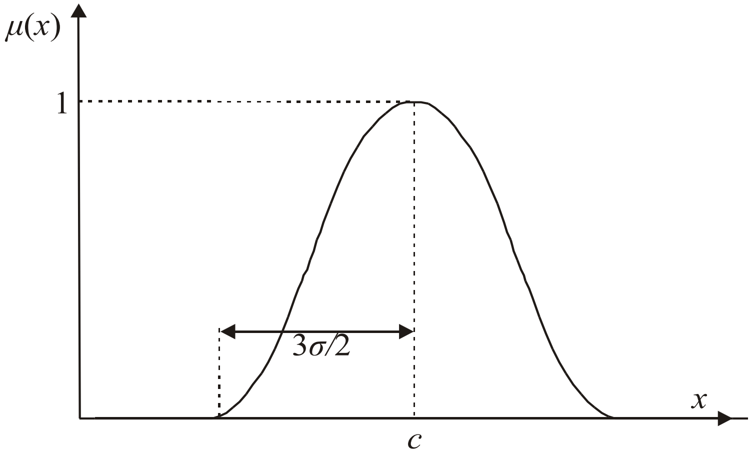
Ασαφές Σύνολο



(α)



(β)



(γ)

Τριγωνικό

$$\mu(x)=\begin{cases} 0 & ,\ \alpha\nu\ x<a \\ \dfrac{x-a}{b-a} & ,\ \alpha\nu\ a\leq x<b \\ -\dfrac{x-c}{c-b} & ,\ \alpha\nu\ b\leq x<c \\ 0 & ,\ \alpha\nu\ x\geq c \end{cases}$$

Τραπεζοειδές

$$\mu(x)=\begin{cases} 0 & ,\ \alpha\nu\ x<a \\ \dfrac{x-a}{b-a} & ,\ \alpha\nu\ a\leq x<b \\ 1 & ,\ \alpha\nu\ b\leq x<c \\ -\dfrac{x-d}{d-c} & ,\ \alpha\nu\ c\leq x<d \\ 0 & ,\ \alpha\nu\ x\geq d \end{cases}$$

Gaussian

$$\mu(x)=\exp[-\frac{(c-x)^2}{\sigma^2}]$$

ΚΑΛΟ ΑΠΟΓΕΥΜΑ