# Trabajo T5

## Ejercicio 1.1

### Algoritmo batch

Entrada: Topología, pesos iniciales $\theta^l_{i\,j,}$ $1 \leq l \leq L$, $1 \leq i \leq M_l$, $0 \leq j \leq M_{l-1}$, factor de aprendizaje $\rho$, condiciones de convergencia, N datos de entrenamiento S, momentum $0 \leq v < 1$.

Salidas: Pesos de las conexiones que minimizan el error cuadrático medio de S.

Mientras no se cumplan las condiciones de convergencia

> Para $1 \leq l \leq L$, $1 \leq i \leq M_l$, $0 \leq j \leq M_{l-1}$, inicializar $\Delta\theta^l_{ij} = 0$
>
> Para cada muestra de entrenamiento $(x, t) \in S$
>
> > Desde la capa de entrada a la de salida ($l = 0, \ldots, L$):
> >
> > > Para $1 \leq i \leq M_l$ si $l = 0$ entonces $s^0_i = x_i$ sino calcular $\phi^l_i$ y $s^l_i = g(\phi^l_i)$
> >
> > Desde la capa de salida a la de entrada ($l = L, \ldots, 1$),
> >
> > > Para cada nodo ($1 \leq i \leq M_l$)
> > >
> > > > Calcular $\delta^l_i$ = si ($l$ == L) then $g'(\phi^l_i)\,(t_{ni} - s^L_i)$ , else $g'(\phi^l_i)\,(\sum_r \delta_r^{l+1}\ \theta_{ri}^{l+1})$
> > > >
> > > > Para cada peso $\theta^l_{ij}$ $(0 \leq j \leq M_{l-1})$ calcular: $\Delta\theta^l_{ij} = v\Delta\theta^l_{ij} + \rho\ \delta^l_i\ s^{l-1}_j$
>
> Para $1 \leq l \leq L$, $1 \leq i \leq M_l$ , $0 \leq j \leq M_{l-1}$, actualizar pesos: $\theta^l_{ij} = \theta^l_{ij} + (1/N)\ \Delta\ \theta^l_{ij}$

## Algoritmo incremental

Entrada: Topología, pesos iniciales $\theta^{l}_{i\,j}$, $1 \le l \le L$, $1 \le i \le M_l$, $0 \le j \le M_{l-1}$, factor de aprendizaje $\rho$, condiciones de convergencia, N datos de entrenamiento S, momentum $0 \le v < 1$.

Salidas: Pesos de las conexiones que minimizan el error cuadrático medio de S.

Mientras no se cumplan las condiciones de convergencia

Para $1 \le l \le L$, $1 \le i \le M_l$, $0 \le j \le M_{l-1}$, inicializar $\Delta\theta^{l}_{ij} = 0$

Para cada muestra de entrenamiento $(x, t) \in S$

Desde la capa de entrada a la de salida ($l = 0, \ldots, L$):

Para $1 \le i \le M_l$ si $l = 0$ entonces $s^{0}_{i} = x_i$ sino calcular $\phi^{l}_{i}$ y $s^{l}_{i} = g(\phi^{l}_{i})$

Desde la capa de salida a la de entrada ($l = L, \ldots, 1$),

Para cada nodo ($1 \le i \le M_l$)

Calcular $\delta^{l}_{i}$ = si ($l == L$) then $g'(\phi^{l}_{i})\,(t_{ni} - s^{L}_{i})$ , else $g'(\phi^{l}_{i})\,(\sum_r \delta_r^{l+1}\ \theta_{ri}^{l+1})$

Para cada peso $\theta^{l}_{ij}$ ($0 \le j \le M_{l-1}$) calcular: $\Delta\theta^{l}_{ij} = \rho\, \delta^{l}_{i}\, s^{l-1}_{j}$

Para $1 \le l \le L$, $1 \le i \le M_l$ , $0 \le j \le M_{l-1}$, actualizar pesos: $\theta^{l}_{ij} = v\theta^{l}_{ij} + (1/N)\, \Delta\, \theta^{l}_{ij}$

# Ejercicio 1.2

## Algoritmo batch

Entrada: Topología, pesos iniciales $\theta^l_{ij}$, $1 \le l \le L$, $1 \le i \le M_l$, $0 \le j \le M_{l-1}$, factor de aprendizaje ρ, condiciones de convergencia, N datos de entrenamiento S, factor de regularización λ.

Salidas: Pesos de las conexiones que minimizan el error cuadrático medio de S.

Mientras no se cumplan las condiciones de convergencia

> Para $1 \le l \le L$, $1 \le i \le M_l$, $0 \le j \le M_{l-1}$, inicializar $\Delta\theta^l_{ij} = 0$
>
> Para cada muestra de entrenamiento $(x, t) \in S$
>
> > Desde la capa de entrada a la de salida ($l = 0, \ldots, L$):
> >
> > > Para $1 \le i \le M_l$ si $l = 0$ entonces $s^0_i = x_i$ sino calcular $\phi^l_i$ y $s^l_i = g(\phi^l_i)$
> >
> > Desde la capa de salida a la de entrada ($l = L, \ldots, 1$),
> >
> > > Para cada nodo ($1 \le i \le M_l$)
> > >
> > > > Calcular $\delta^l_i$ = si ($l$ == L) then $g'(\phi^l_i)(t_{ni} - s^L_i)$ , else $g'(\phi^l_i)(\sum_r \delta_r^{l+1}\ \theta_{ri}^{l+1})$
> > > >
> > > > Para cada peso $\theta^l_{ij}$ ($0 \le j \le M_{l-1}$) calcular: $\Delta\theta^l_{ij} = \Delta\theta^l_{ij} - \rho\,\lambda\,\delta^l_i\, s^{l-1}_j$
>
> Para $1 \le l \le L$, $1 \le i \le M_l$ , $0 \le j \le M_{l-1}$, actualizar pesos: $\theta^l_{ij} = \theta^l_{ij} + (1/N)\,\Delta\,\theta^l_{ij}$

## Algoritmo incremental

Entrada: Topología, pesos iniciales $\theta^l_{i\,j}$, $1 \leq l \leq L$, $1 \leq i \leq M_l$, $0 \leq j \leq M_{l-1}$, factor de aprendizaje ρ, condiciones de convergencia, N datos de entrenamiento S, factor de regularización λ.

Salidas: Pesos de las conexiones que minimizan el error cuadrático medio de S.

Mientras no se cumplan las condiciones de convergencia

Para $1 \leq l \leq L$, $1 \leq i \leq M_l$, $0 \leq j \leq M_{l-1}$, inicializar $\Delta\theta^l_{ij} = 0$

Para cada muestra de entrenamiento $(x, t) \in S$

Desde la capa de entrada a la de salida ($l = 0, \ldots, L$):

Para $1 \leq i \leq M_l$ si $l = 0$ entonces $s^0_i = x_i$ sino calcular $\phi^l_i$ y $s^l_i = g(\phi^l_i)$

Desde la capa de salida a la de entrada ($l = L, \ldots, 1$),

Para cada nodo ($1 \leq i \leq M_l$)

Calcular $\delta^l_i$ = si ($l == L$) then $g'(\phi^l_i)\,(t_{ni} - s^L_i)$ , else $g'(\phi^l_i)\,(\sum_r \delta_r^{l+1}\ \theta_{ri}^{l+1})$

Para cada peso $\theta^l_{ij}$ ($0 \leq j \leq M_{l-1}$) calcular: $\Delta\theta^l_{ij} = \rho\ \delta^l_i\ s^{l-1}_j - p\ \lambda\Delta\theta^l_{ij}$

Para $1 \leq l \leq L$, $1 \leq i \leq M_l$ , $0 \leq j \leq M_{l-1}$, actualizar pesos: $\theta^l_{ij} = \theta^l_{ij} + (1/N)\ \Delta\ \theta^l_{ij}$

## Ejercicio 2

$$\Delta\theta_{ij}^{l} = \frac{\partial \mathrm{qs}(\theta)}{\partial \theta_{ij}^{l}} = \mathrm{p}\,\frac{t_i}{s_i^2} f'\left(z_i^l\right) s_j^{l-1} = \mathrm{p}\delta^l s_j^{l-1}$$