



eset el procedimiento que se usa par el forward
y para el back ward se multiplican los errores por la transpuesta de los pesos para propagar el gradiente hacia atrás luego se calculan las derivadas para ajustar los pesos multiplicando las activaciones previas por esos gradientes
como se usa full matrices lo que se paralelisa es justo las operaciones basicas de estas

```

inicializar_aleatorio()
multiplicar()
sumar()
restar()
transponer()
hadamard()
aplicar_funcion()
multiplicar_escalar()
llenar()
mezclar_datos()
propagar_hacia_atras()
calcular_error_epoca()
calcular_precision()
entrenar_lote()

```

