



Aprendizaje Automático Profundo (Deep Learning)

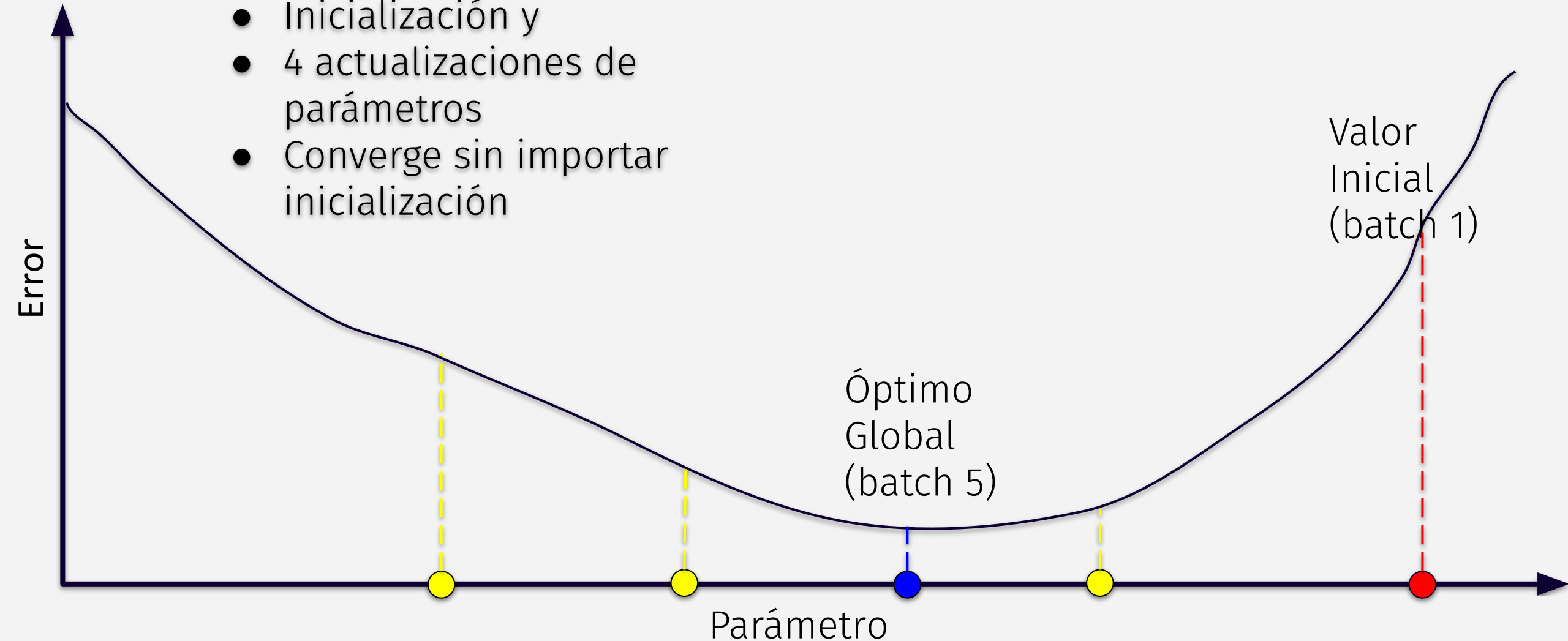
Dr. Facundo Quiroga - Dr. Franco Ronchetti



Inicialización de Parámetros

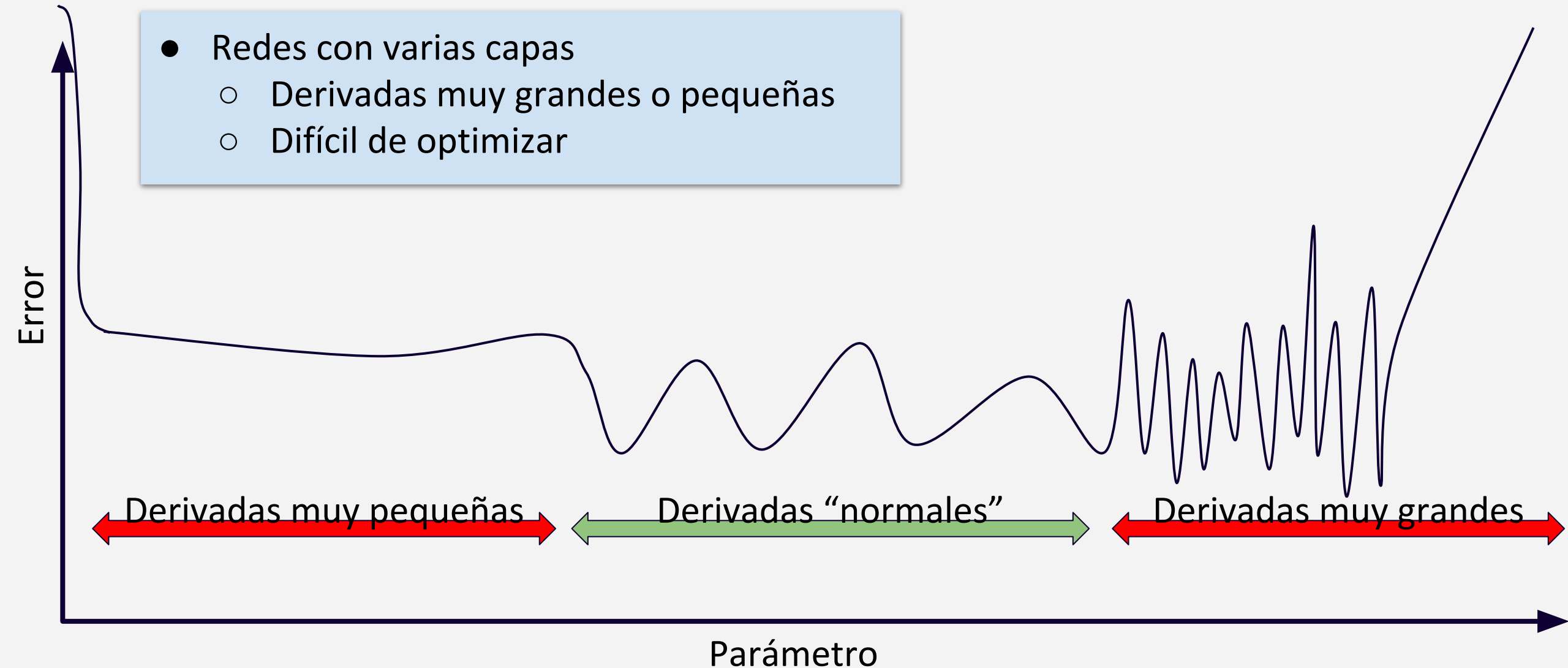
Valor inicial en optimización convexa

- Inicialización y
- 4 actualizaciones de parámetros
- Converge sin importar inicialización



Valor inicial en optimización no convexa

- Redes con varias capas
 - Derivadas muy grandes o pequeñas
 - Difícil de optimizar



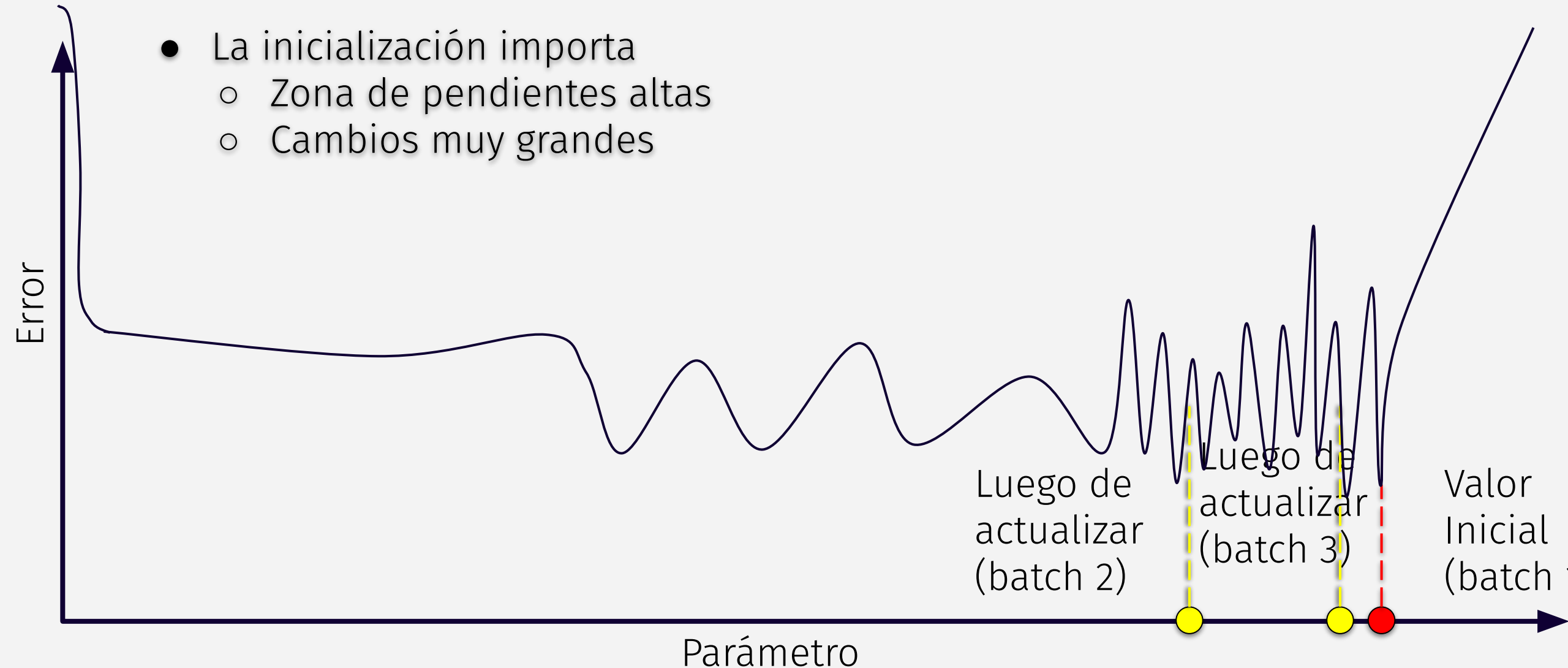
Inicialización de parámetros

- La inicialización importa
 - Zona de gradientes bajos
 - Cambios muy chicos



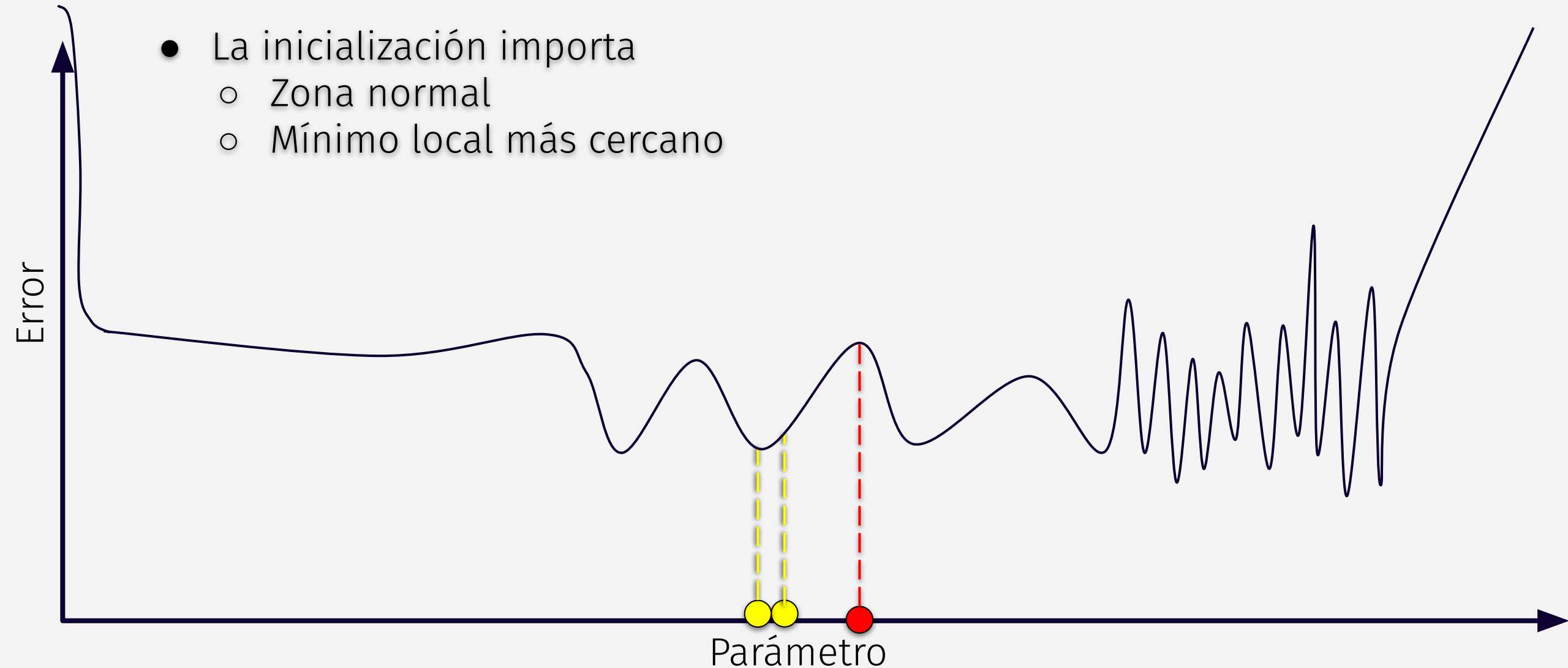
Inicialización de parámetros

- La inicialización importa
 - Zona de pendientes altas
 - Cambios muy grandes



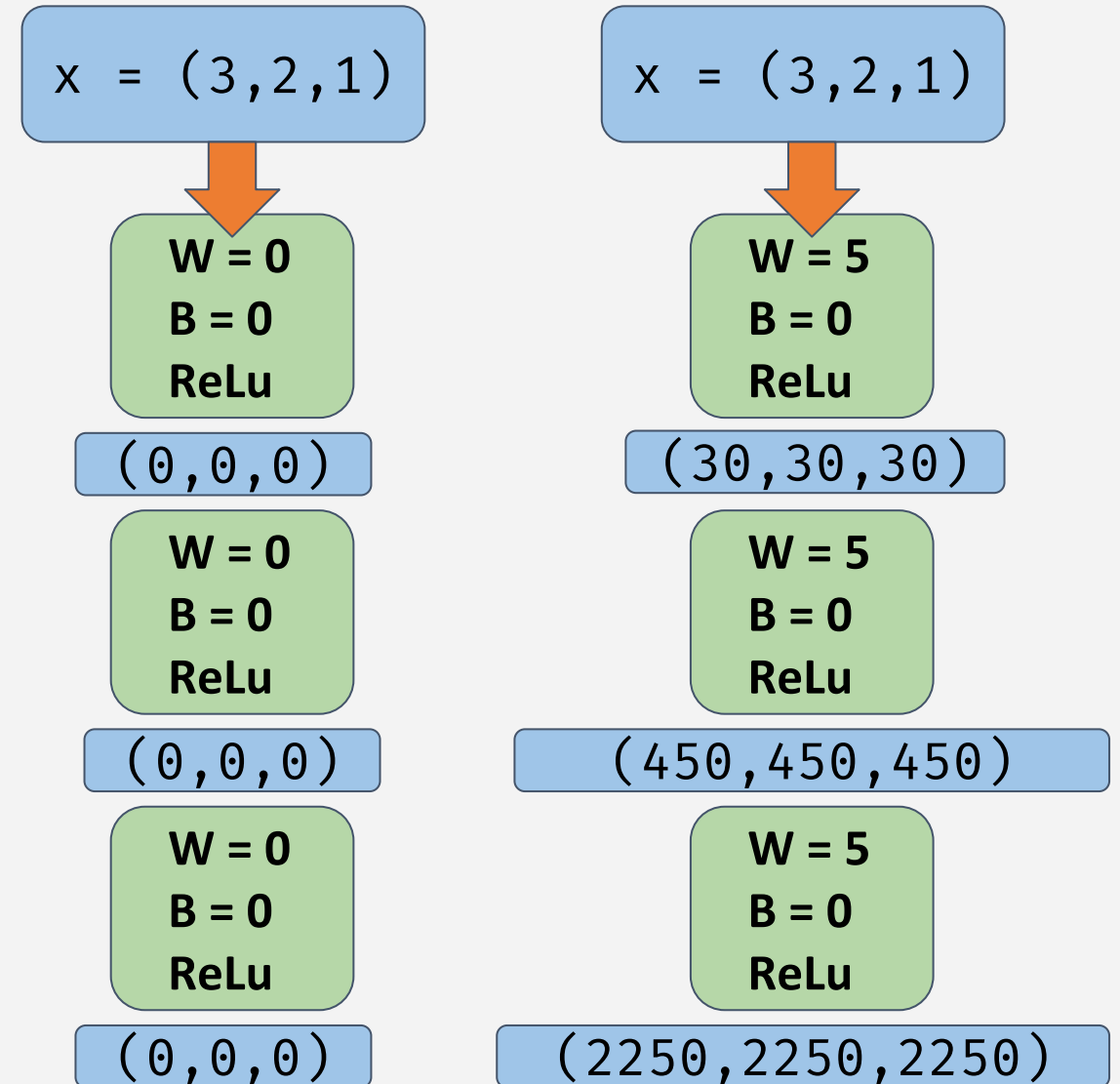
Inicialización de parámetros

- La inicialización importa
 - Zona normal
 - Mínimo local más cercano



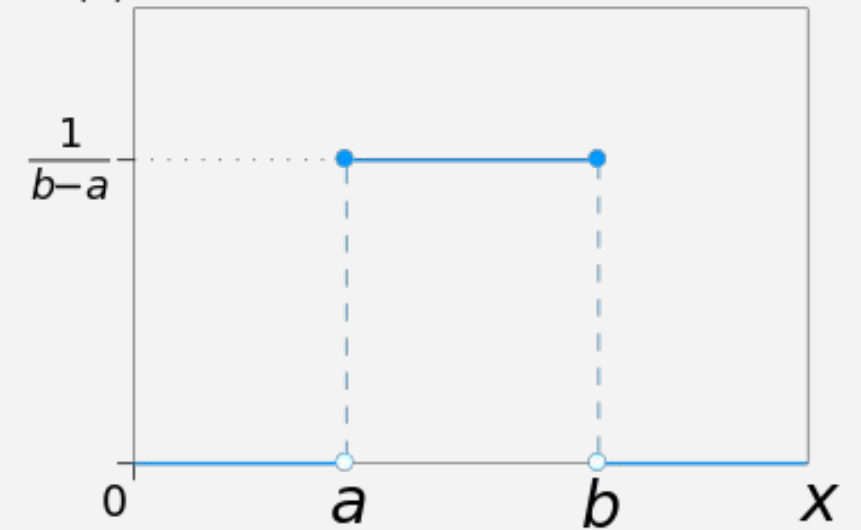
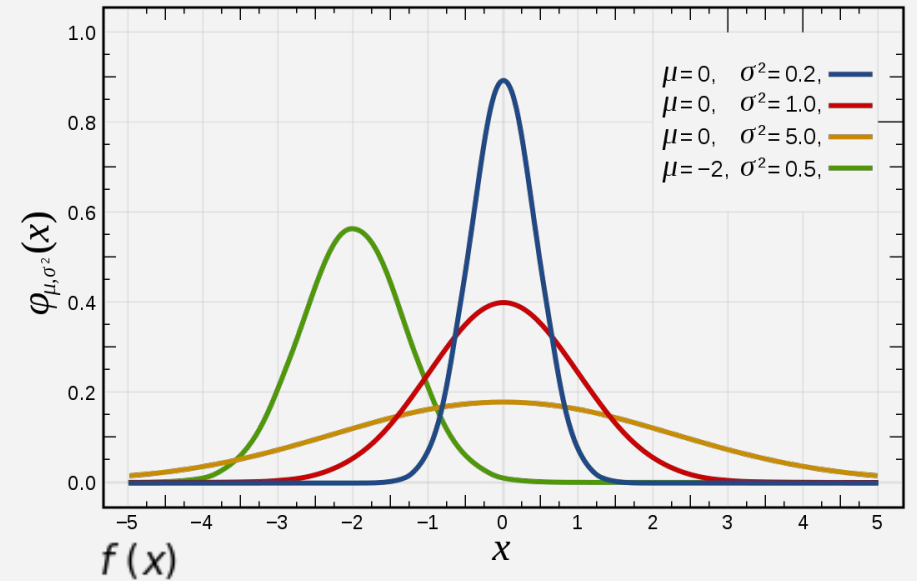
Inicialización con valores fijos

- Generalmente 0
 - Funciona para sesgos B
 - No funciona para matrices de pesos W o filtros convolucionales
 - Predice todo 0
 - Las derivadas son 0
 - No aprende
- Otro valor fijo (5, por ejemplo)
 - Todas las neuronas predicen lo mismo
 - Simetría



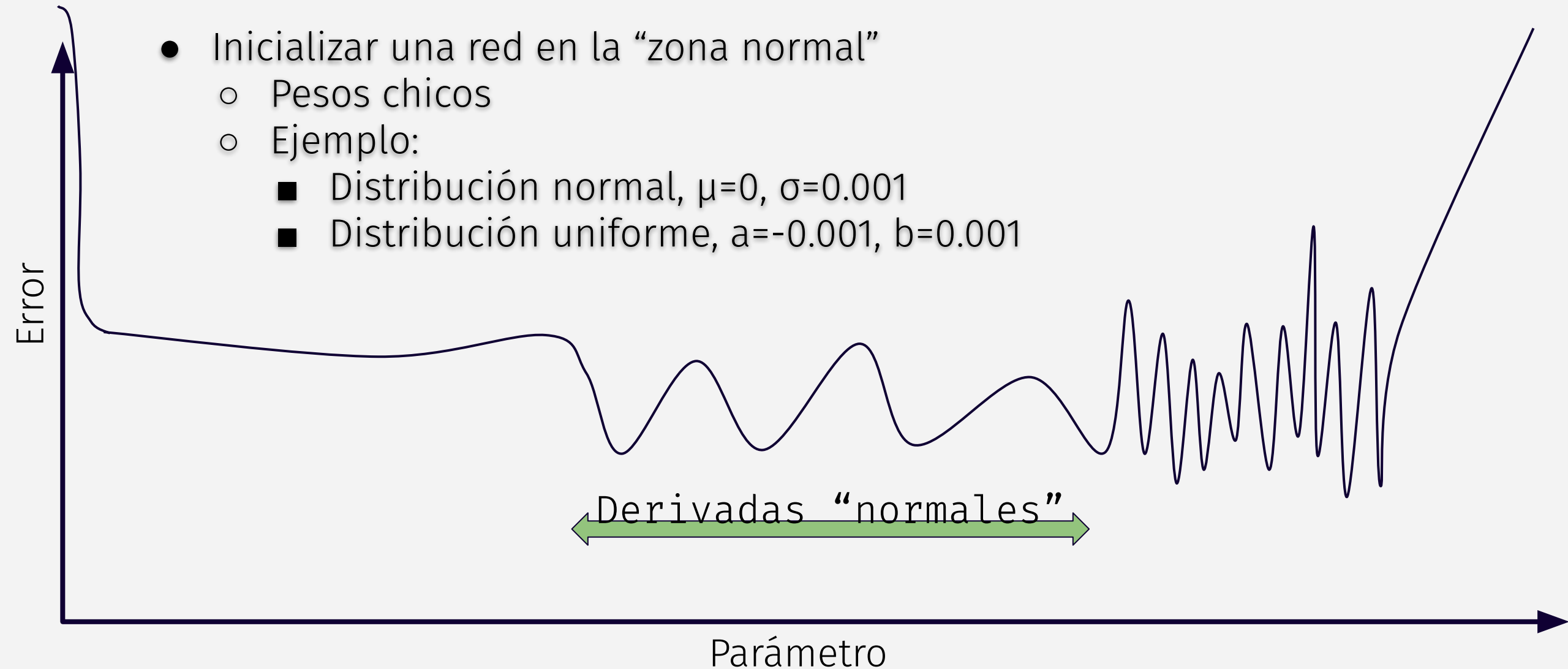
Inicialización con valores aleatorios

- Distribución normal o gaussiana
 - Elegir μ y σ (hiperparámetros)
- Distribución uniforme
 - Elegir rango del intervalo: a y b (hiperparámetros)
- ¿El entrenamiento no funcionó?
 - Repetir inicialización
 - Volver a entrenar



Escala de valores aleatorios

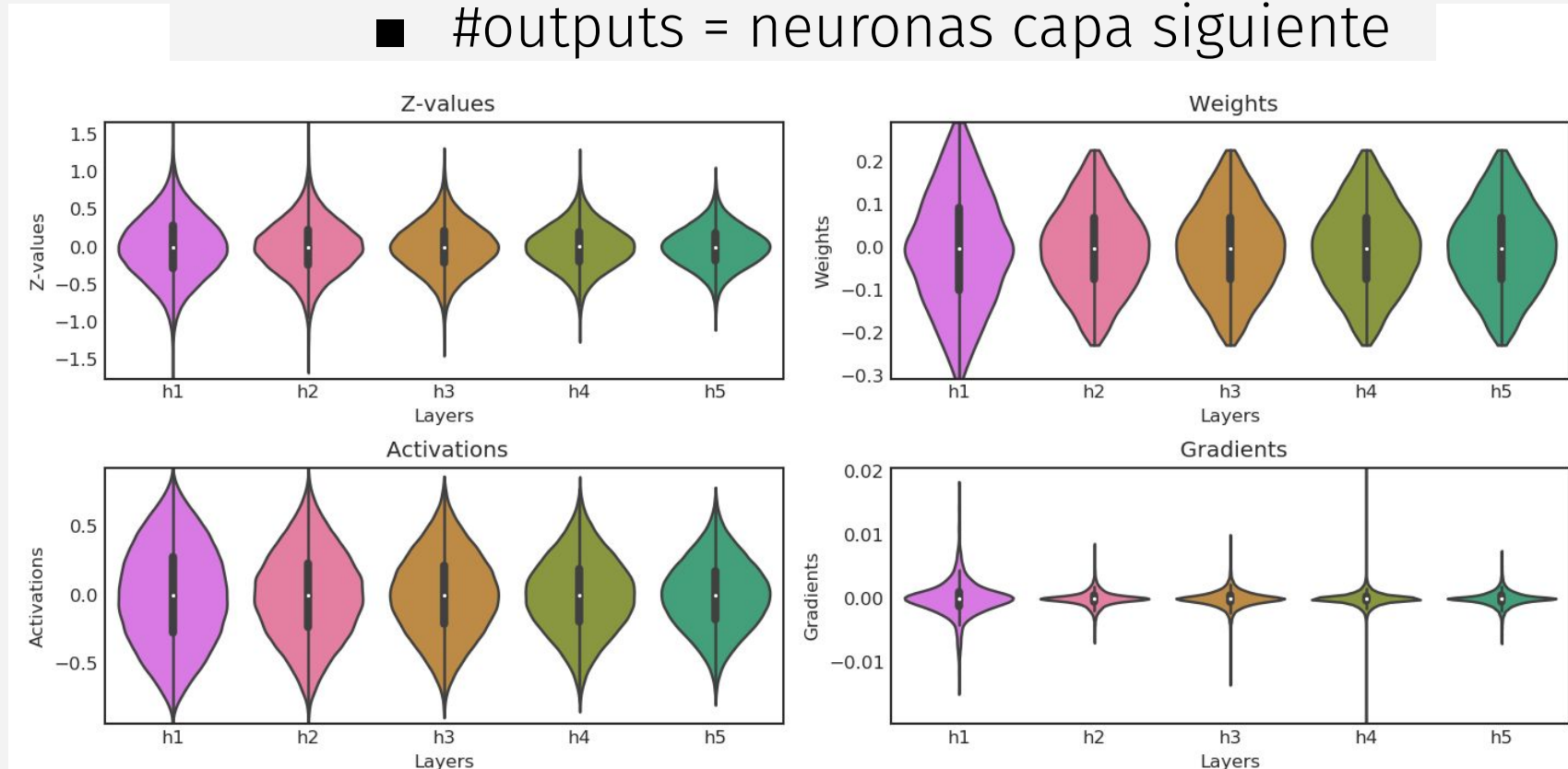
- Inicializar una red en la “zona normal”
 - Pesos chicos
 - Ejemplo:
 - Distribución normal, $\mu=0$, $\sigma=0.001$
 - Distribución uniforme, $a=-0.001$, $b=0.001$



Inicializador Glorot para TanH

- Inicializadores avanzados
- Objetivo: magnitudes de los valores intermedios *estables*
 - estable:
 - No tienden a 0
 - Ni a valores muy grandes
 - => Derivadas estables
- Inic. más importantes
 - Glorot: para capas con TanH
 - He: para capas con ReLu

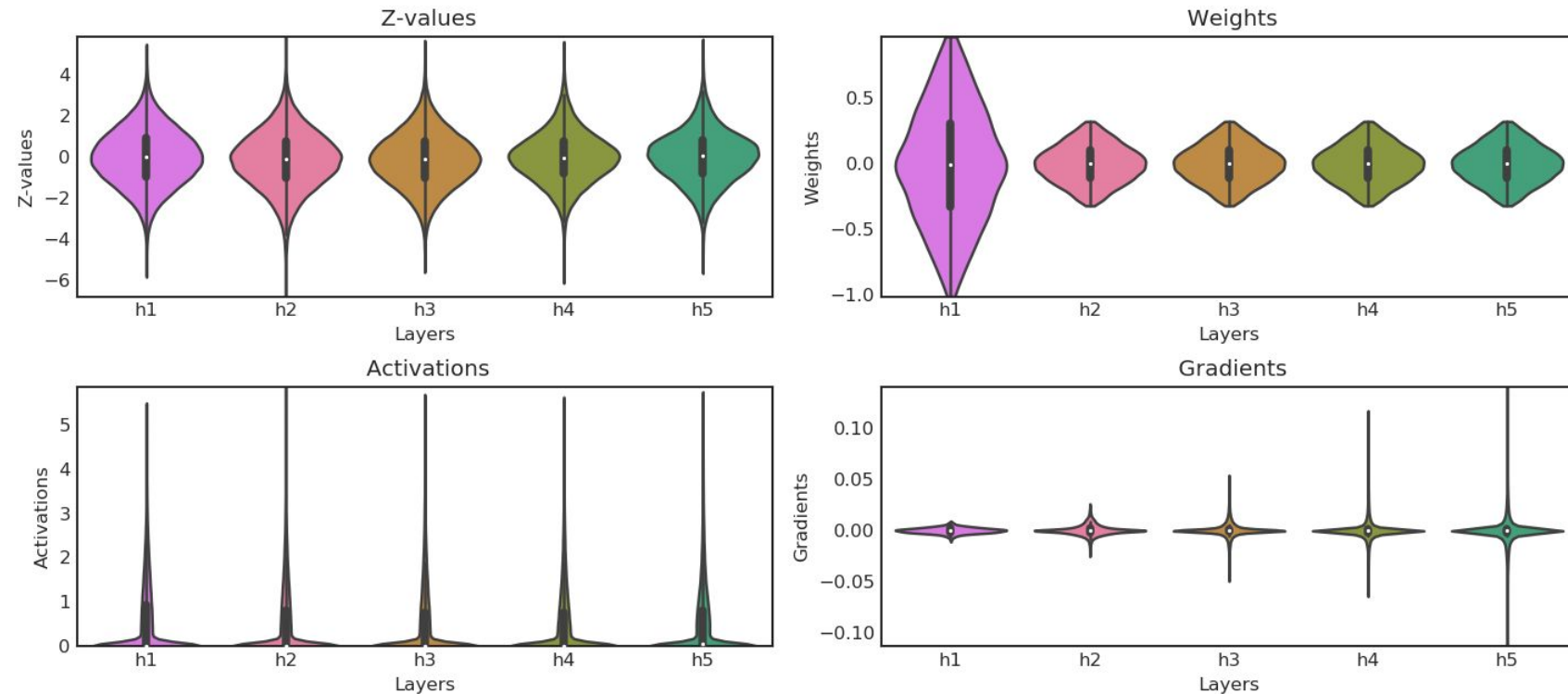
- Glorot: Normal con media = 0 y
 - Varianza = $2 / (\text{\#inputs} + \text{\#outputs})$
 - \#inputs = neuronas capa anterior
 - \#outputs = neuronas capa siguiente



Inicializador He para ReLU ([más info](#))

- Inicializadores avanzados
- Objetivo: magnitudes de los valores intermedios *estables*
 - estable:
 - No tienden a 0
 - Ni a valores muy grandes
 - => Derivadas estables
- Inic. más importantes
 - Glorot: para capas con TanH
 - He: para capas con ReLu

- He: Normal con media = 0 y
 - Varianza = $2 / (\#inputs)$
 - $\#inputs$ = neuronas capa anterior



Inicialización de parámetros - Keras

- Inicializadores por defecto en keras
 - Clase Dense
 - **kernel_initializer**: inicializador de matriz de pesos (w)
 - **bias_initializer**: inicializador de vector de sesgo (b)
 - Clase Conv2D
 - **kernel_initializer**: inicializador de pesos de filtros convolucionales (w)
 - **bias_initializer**: inicializador de vector de sesgo (b)
 - Por defecto **'glorot_uniform'** en pesos
 - Cambiar a **'he_uniform'** o **'he_normal'** si usás ReLu

```
Dense( ... ,kernel_initializer='glorot_uniform', bias_initializer='zeros', .. )
```

```
Conv2D( ... ,kernel_initializer='glorot_uniform', bias_initializer='zeros', .. )
```