

# Tarea 1. Algoritmos de descenso estocástico

Aprendizaje de Máquina I

February 3, 2021

**Descripción.** Implementar versiones estocásticas de algoritmos de descenso de gradiente:

1. Descenso de gradiente estocástico.
2. Descenso de gradiente estocástico acelerado de tipo Nesterov.
3. AdaDelta
4. ADAM
5. NADAM

Para resolver el problema de regresión:

$$\min_{\alpha, \mu} F(\alpha, \mu) = \frac{1}{2} \|\Phi\alpha - y\|_2^2 \quad (1)$$

donde

$$\Phi = [\phi_1, \phi_2, \dots, \phi_n]$$

con

$$\phi_{ij} = \exp\left(-\frac{1}{2\sigma^2}(j - \mu_i)^2\right);$$

$j$  : representa las posiciones donde se tienen muestras en una serie de tiempo;

$y$  : es un vector de dimensión  $n$  que representa los valores de la serie de tiempo;

$\alpha$  : es un vector de coeficientes de dimensión  $m$  cuyas entradas representan la contribución de cada una de las  $j$ -ésimas funciones base  $\phi_j$  para aproximar óptimamente  $y$ ;

$\sigma$  : es una constante que determina el ancho de cada función base.

$\phi_i$  : es el  $i$ -ésimo vector columna (función radial base) cuya contribución para ajustar los datos  $y$  esta dada por el coeficiente  $\alpha_i$

## Estrategia de solución

Dado un vector de medias iniciales  $\mu_0$  y  $t = 0$ , para resolver (1) usar alternar hasta convergencia:

1. Resolver (aproximadamente)

$$\alpha_{t+1} = \arg \min_{\alpha} F(\alpha, \mu_t)$$

con  $\mu_t$  fija.

2. Resolver aproximadamente

$$\mu_{t+1} = \arg \min_{\mu} F(\alpha_{t+1}, \mu)$$

con  $\alpha_{t+1}$  fija.

3. Hacer  $t=t+1$

### Entrega de la tarea

La tarea se entrega como el fuente del notebook de jupyterlab (.pynb) con la ltima ejecucin. Usar  $n \geq 1000$ ,  $m \geq 15$ .

**Enviar la tarea** a aprendizaje.maquina@cimat.mx. Con asunto: “Tarea *número\_de\_tarea*. grupo *nombre\_del\_curso\_inscrito*”. Ejemplo: Tarea 1. grupo Aprendizaje Automático I

**Fecha de entrega:** 9 de febrero 2021 a las 12pm (límite).

**Penalización por retraso:** la calificación de la tarea se multiplicará por  $0.9^n$  donde  $n \geq 0$  son los días de retraso.

### Material de apoyo

[https://www.cimat.mx/~mriviera/cursos/optimizacion/descenso\\_grad\\_estocastico/descenso\\_grad\\_estocastico.html](https://www.cimat.mx/~mriviera/cursos/optimizacion/descenso_grad_estocastico/descenso_grad_estocastico.html)