Optimización

Primer Cuatrimestre 2025

Posibles temas de TP final

Ejercicio 1 (Restricciones dadas por ODEs). Deducción del Principio del Máximo de Pontryagin a partir de multiplicadores de Lagrange y Cálculo Variacional. Aplicación a un problema de control óptimo a elección (ej. aterrizaje del cohete).

Ejercicio 2 (Restricciones dadas por PDEs). Deducción del método adjunto, a partir de Multiplicadores de Lagrange, para un problema de optimización donde las restricciones están dadas por la solución de una ecuación en derivadas parciales. Por ejemplo, el problema de encontrar la velocidad del sonido $c(x)$ de un medio que focaliza una onda incidente dada en un punto deseado, o el problema de encontrar el valor de $c(x)$ que mejor ajusta los datos de un sistema de tomografía.

Ejercicio 3 (PCA Robusta) El análisis de componentes principales (PCA) tradicional es muy sensible frente a outliers. Existen alternativas robustas (RPCA) basadas en descomposición de matrices $A = L + S$, donde L es una matriz de rango bajo (Low-rank) y S una matriz rala (Sparse), u otras variantes. La RPCA puede aplicarse, bajo ciertas condiciones, a problemas como remover el fondo de un video, separar la voz del acompañamiento en una canción, o para eliminar

Un artículo fundamental que aborda este problema es *"Robust Principal Component Analysis?"* de Emmanuel J. Cand'es,, Xiaodong Li, , Yi Ma, and John Wright4, arXiv, 2009.

Se pide implementar el método del paper y estudiar la teoría subyacente.

Ejercicio 4 (Problema de Netflix). El problema de reconstrucción de una matriz rala a partir de observaciones parciales puede resolverse eficientemente mediante métodos proximales. Por ejemplo, en el artículo *"A singular value thresholding algorithm for matrix completion"* de Cai, Jian-Feng and Candès, Emmanuel J and Shen, Zuowei, SIAM Journal on optimization (2010), se propone uno de dichos métodos.

Se pide implementar el método del paper y estudiar la teoría subyacente.

Ejercicio 5 (Redes convolucionales) Modifique el trabajo práctico número 1 para incorporar las siguientes mejoras:

(a) Redes convolucionales (cambia la arquitectura de la red).

(b) La elección del paso mediante momentum o métodos adaptivos Adam, RMSProp.

Ejercicio 6 (Deep learning). Uno de los problemas que trabó el desarrollo de redes neuronales profundas durante muchos años es el problema de los gradientes que se anulan o explotan para secuencias largas.

(a) Muestre que el gradiente puede escribirse como una productoria (...)

(b) Implemente una solución para este problema (ResNets)

Ejercicio 7 (Modelos Generativos) El Variational Autoencoder se basa en el entrenamiento de dos redes neuronales: una que comprime (o codifica) a un espacio latente de baja dimensión ($q_\phi(z|x)$), y otra que decodifica ($p_\theta(x|z)$. Para ello, se minimiza la siguiente función:

$$\mathcal{L}(\theta, \phi; x) = -\mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] + D_{KL}(q_\phi(z|x)||p(z))$$

donde $p(z)$ es una distribución a-priori sobre el espacio latente, y D_{KL} la divergencia de Kullback-Leibler.

(a) Aplicando la reparametrización de la variable latente $z = \mu_\phi(X) + \sigma_\phi(x) \cdot \epsilon$, donde $\epsilon \sim N(0, 1)$ es independiente de los parámetros del modelo, aplique el método de back-propagation para minimizar $\mathcal{L}$.

(b) Modifique el código del trabajo práctico número 1 para entrenar un VAE sobre el dataset MNIST.

Ejercicio 8 (Gradient Boosting)

(a) Entender cómo AdaBoost, inicialmente concebido como una heurística ingeniosa, puede ser rigurosamente demostrado como un caso específico de descenso de gradiente funcional al minimizar una función de pérdida exponencial.

(b) Entender cómo este nuevo marco teórico (Gradient Boosting) permite permite su adaptabilidad a diversos problemas de clasificación y regresión.

Ejercicio 9 (Entrenamiento de Support Vector Machines (SVM)). El problema de entrenar un SVM no-lineal se formula utilizando el dual Lagrangiano y aplicando una generalización llamada "kernel-trick", lo que da lugar a un problema de optimización cuadrática convexo. El algoritmo más eficiente para este problema se conoce como Sequential Minimal Optimization (SMO), que se basa en optimizar un par de multiplicadores de Lagrange a la vez, dejando los demás fijos. Se pide entender esta teoría e implementar el algoritmo SMO desde cero, verificando su convergencia.

Ejercicio 10 (Optimización combinatoria). Algunos temas de optimización combinatoria que se benefician de enfoques basados en optimización convexa y/o dualidad Lagrangiana:

(a) Relajación Lagrangiana para el problema de la mochila. La resolución de un problema dual se utiliza para obtener cotas dentro de un esquema de Branch-and-Bound.

(b) Branch-and-Cut mediante cortes de Benders. Estos cortes se derivan del problema dual Lagrangiano.

(c) Branch-and-Price. Es una técnica para resolver problemas con un gran número de variables. Se basa en una reformulación de Dantzig-Wolfe, que tiene una profunda conexión con la dualidad Lagrangiana.

(d) Algoritmo de Goemans-Williamson: Es una relajación convexa del problema de Max-Cut, que maneja las variables binarias a través de una matriz semidefinida positiva. El problema resultante es de tipo SDP.