**FUNDAMENTO DE OPTIMIZACIÓN 1S-2024**
**RESUMEN DE CLASES**

Clase 1 (18/04/2024) - Introducción al curso y repaso

Para introducir el curso comenzamos con una presentación hecha por el prof. Marcelo Fiori donde se muestran diversos ejemplos de optimización en problemas de ingeniería (para ver lo slides puede hacer click AQUI).

Luego comenzamos con un repaso general de algunos conceptos que se ven mayoritariamente en Cálculo 2 como ser:

(1) El diferencial $df_p$, la matriz Jacobiana $\mathrm{Jf}(p)$, el gradiente $\nabla f(p)$, derivadas parciales $\frac{\partial f}{\partial x_i}(p)$ y direccionales $\frac{\partial f}{\partial v}(p)$. Recordamos algunas importantes propiedades de las derivadas direccionales:
   (a) $\frac{\partial f}{\partial v}(p) > \frac{\partial f}{\partial w}(p)$ entonces para algún $\epsilon > 0$ se cumple $f(p + vt) > f(p + wt)$ para todo $t \in (0, \epsilon)$.
   (b) Si $\nabla f(p) \neq 0$, $v_0 = \frac{\nabla f(p)}{||\nabla f(p)||}$ entonces $\frac{\partial f}{\partial v_0}(p) > \frac{\partial f}{\partial v}(p)$, $\forall v : ||v|| = 1$ (se lee: el gradiente es la dirección de mayor crecimiento).
   (c) Si $\nabla f(p) \neq 0$, $v_0 = -\frac{\nabla f(p)}{||\nabla f(p)||}$ entonces $\frac{\partial f}{\partial v_0}(p) < \frac{\partial f}{\partial v}(p)$, $\forall v : ||v|| = 1$ (se lee: el gradiente opuesto es la dirección de mayor decrecimiento).
   (d) El signo de $\frac{\partial f}{\partial v}(p)$ indica crecimiento/decrecimiento (si es positivo/negativo entonces la función en $p$ crece/decrece en la dirección $v$).

(2) Condición necesaria de optimalidad: $\nabla f(x^*) = 0$ para $x^*$ mínimo local.

(3) Ejemplo: Minimizar $f^2_{\mathbb{R}} \to R$ dada por $f(x, y) = x^2 + y^2$
   - sin restricciones (aprovechamos para repasar conceptos como curvas de nivel y vimos que el gradiente es perpedicular a las curvas de nivel)
   - con una restricción del tipo $g(x, y) = 0$ (método de los multiplicadores de Lagrange).

Comentario: El método de los multiplicadores de Lagrange sirve para minimizar una función $f(x)$ sujeto a la restricción del tipo $g(x) = 0$ (donde $f, g : \mathbb{R}^n \to \mathbb{R}$ son de clase $C^1$). Se considera la función Lagrangiana (o Lagrangiano) $F(x, \lambda) = f(x) - \lambda g(x)$. El Teorema de los multiplicadores de Lagrange dice que si $f$ tiene mínimo en $x = x^*$ y $\nabla g(x^*) \neq 0$ entonces existe un $\lambda^*$ que verifica $\nabla F(x^*, \lambda^*) = 0$. Entonces la idea es encontrar todos los $(x, \lambda)$ para los cuales $\nabla F = 0$ y chequear esos posibles $x$ cual minimiza $f$ (debemos estar seguros que $f$ tenga mínimo absoluto en la región considerada sino solo vamos a encontrar un mínimo relativo).

Lectura introductoria recomendada sobre el tema:

- Unas notas (en las páginas 1-3 ya da una idea básica del método)
- Wikipedia

## Clase 2 (22/04/2024) - Repaso y más ejemplos

Repasamos algunas propiedad de la Jacobiana (similares a las propiedades que cumple la derivada para funciones de una variable): $J(f+g)(p) = Jf(p) + Jg(p)$, $J(fg)(p) = f(p)Jg(p) + Jf(p)g(p)$ y $J(f \circ g)(p) = Jf(g(p)) \cdot Jg(p)$ (cuando las operaciones tienen sentido y las funciones son diferenciables en todos los lados donde se necesita).

Norma y productos internos en espacios vectoriales. Como ejemplos importantes vimos:

- El producto interno usual de $\mathbb{R}^n$: $\langle x, y \rangle = x^t y$, que da lugar a la norma euclídea (denotada por $||x||_2$ o simplemente por $||x||$ cuando no hay peligro de ambigüedad).
- El producto interno en $\mathbb{R}^{n \times m}$: $\langle A, B \rangle = \mathrm{tr}(A^t \cdot B)$, que da lugar a la norma de Frobenius $||A||_F = \sum_{i,j} A_{i,j}^2$ (juega un papel importante en aplicaciones prácticas).
- La norma espectral: $||A||_2 = \max\{Ax : ||x|| = 1\}$ que cumple la propiedad $||Ax||_2 \leq ||A||_2 \cdot ||x||_2$ (esa propiedad la hace muy útil para probar resultados teóricos). Esta norma también se denote simplemente por $||A||$ si no hay peligro de ambigüedad y la propiedad mencionada se puede escribir como $||Ax|| \leq ||A|| \cdot ||x||$. Es importante observar que $||A||$ es un número real que no depende de $x$.

Luego vimos ejemplos varios de cálculos de gradiente, mostrando dos formas de cálculo:

- Usando la unicidad de la Jacobiana y la propiedad $||Ax|| \leq ||A|| \cdot ||x||$ para acotar el resto.
- Usando la regla de la cadena (quedó como ejercicio completar detalles).

Los ejemplos fueron:

- Cálculo del gradiente de una forma cuadrática $q : \mathbb{R}^n \to \mathbb{R}$, $q(x) = x^t Ax$.
- Cálculo del gradiente de una función $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = ||Ax||^2$.

Para finalizar mostramos una forma de chequeo para el cálculo del gradiente, aproximando las derivadas parciales usando la definición de límite.

Comentario 1: el contenido corresponde al video 1 del 2023, más un poco de contenido extra como multiplicadores de Lagrange, la norma espectral y vimos con un poco más de detalle el cálculo del gradiente usando la unicidad del Jacobiano y la propiedad de la norma espectral, no solo para el ejemplo de mínimos cuadrados sino también para formas cuadráticas. Para el ejemplo de mínimos cuadrados también mostramos como calcular el gradiente usando la regla de la cadena para Jacobianos.

Comentario 2: Para aquellos estudiantes que no pudieron asistir a clases y están utilizando los videos del curso anterior recomiendo escoger la opción 1 del obligatorio 1 para el cual alcanza con lo visto en el video 1 del 2023.