Sistemas de ecuaciones:

Supongamos que tenemos que resolver un sistema de ecuaciones que se puede escribir así:

$$\vec{f}(\vec{x}) = \vec{b}$$

luego

$$\vec{f}(\vec{x}) - \vec{b} = \vec{0} \quad , \quad \vec{x} \text{ es una raíz de la ecuación.}$$

Si $\vec{f}(\vec{x})$ es una función lineal, lo que nos dice el formalismo matricial es que la dimensión de $\vec{x}$ debe ser al menos de la dimensión de $\vec{f}$...

(Despierta muchas inquietudes)

(A) Pensemos en sistemas lineales por ahora.

$$\mathbb{A}\,\mathbb{X} = \mathbb{B} \quad \rightarrow \quad \mathbb{X} = \mathbb{A}^{-1}\,\mathbb{B}$$

Sólo es válido si es un sistema "cuadrado" e invertible

Entonces, ¿qué sucede si no es cuadrado?

$$\mathbb{A}_{m\times n}\,\mathbb{X}_{n\times p} = \mathbb{B}_{m\times p}$$

/ Situación ideal $m \geq n$

ojalá $m \gg n$!

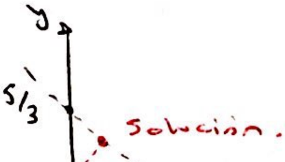
* Pseudo inversa $(m \leq n)$ — No recomendable.

$$\mathbb{X}_{n\times p} = \left(A^{-1}_{m\times n}\right)_{n\times m} \cdot B_{m\times p}$$

La solución más cercana a $(0, 0, 0, \ldots)$

Ejemplo "radical":

$$2x + 3y = 5$$



* Usando la máxima verosimilitud.

Tomemos "$AX = b$"

$A x = \hat{b}$; $\hat{b}$ son las predicciones.

$L(\hat{b}, b_{obs}) :=$ Distancia entre las dos: función de costo

$$L(\hat{b}, b_{obs}) = \frac{1}{N}(\hat{b} - b_{obs})^T (\hat{b} - b_{obs})$$

- Positiva
- Convexa
- Suave
- ¡Tiene un mínimo!

Resultado: $X_{OPT} = (A^T A)^{-1} \cdot A^T b_{obs}$

¡Demostrar!

- ¡Adicionalmente puedo usar un método iterativo para encontrar el optimo! Gradiente decendiente

¿Cómo calculo el r^2?

$$A := \begin{pmatrix} x_1 & y_1 & z_1 & \cdots \\ x_2 & y_2 & z_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{array}{l} \leftarrow \text{Obs } 1 \equiv \vec{x}_1 \\ \leftarrow \text{Obs } 2 \equiv \vec{x}_2 \\ \vdots \end{array}$$

Correlaciones entre variables

$$Cov = \frac{1}{m} A^T A - \mu^T \cdot \mu$$

Con $\mu = (\bar{x} \ \ \bar{y} \ \ \bar{z} \ \ \ldots)$ ← Observaciones medias!

$$r^2 = \frac{\sum_j (\hat{b}_j - \bar{b})^2}{\sum_j (b_j - \bar{b})^2} = \frac{1}{N\sigma_b^2} \sum_j (\hat{b}_j - \bar{b})^2$$

¡Implementemos!

Scanned with CamScanner

Gradiente Descendiente y Newton Rapson.

$$\frac{\partial L}{\partial \vec{x}} = \frac{2}{N}\left(A^T A x - A^T b\right)$$

$$\frac{\partial^2 L}{\partial \vec{x}^2} = \frac{2}{N}\left(A^T A\right)$$ "Definida positiva"

Dado un $\vec{x}_0$, por ejemplo $\vec{0}$, yo puedo usar Newton Rapson para encontrar el siguiente $\vec{x}_1$ que minimiza L ¡Encontrar una raiz!

$L(\hat{b}, b_{obs}) \geq l_{min} \Rightarrow L(\hat{b}, b_{obs}; \vec{x}) - l_{min}$ | tiene al menos 1 raiz

$$\vec{x}_j = \vec{x}_{j-1} - \frac{\partial L}{\partial \vec{x}}\epsilon$$ con ϵ un parámetro pequeño "Tasa de aprendizaje"

Scanned with CamScanner