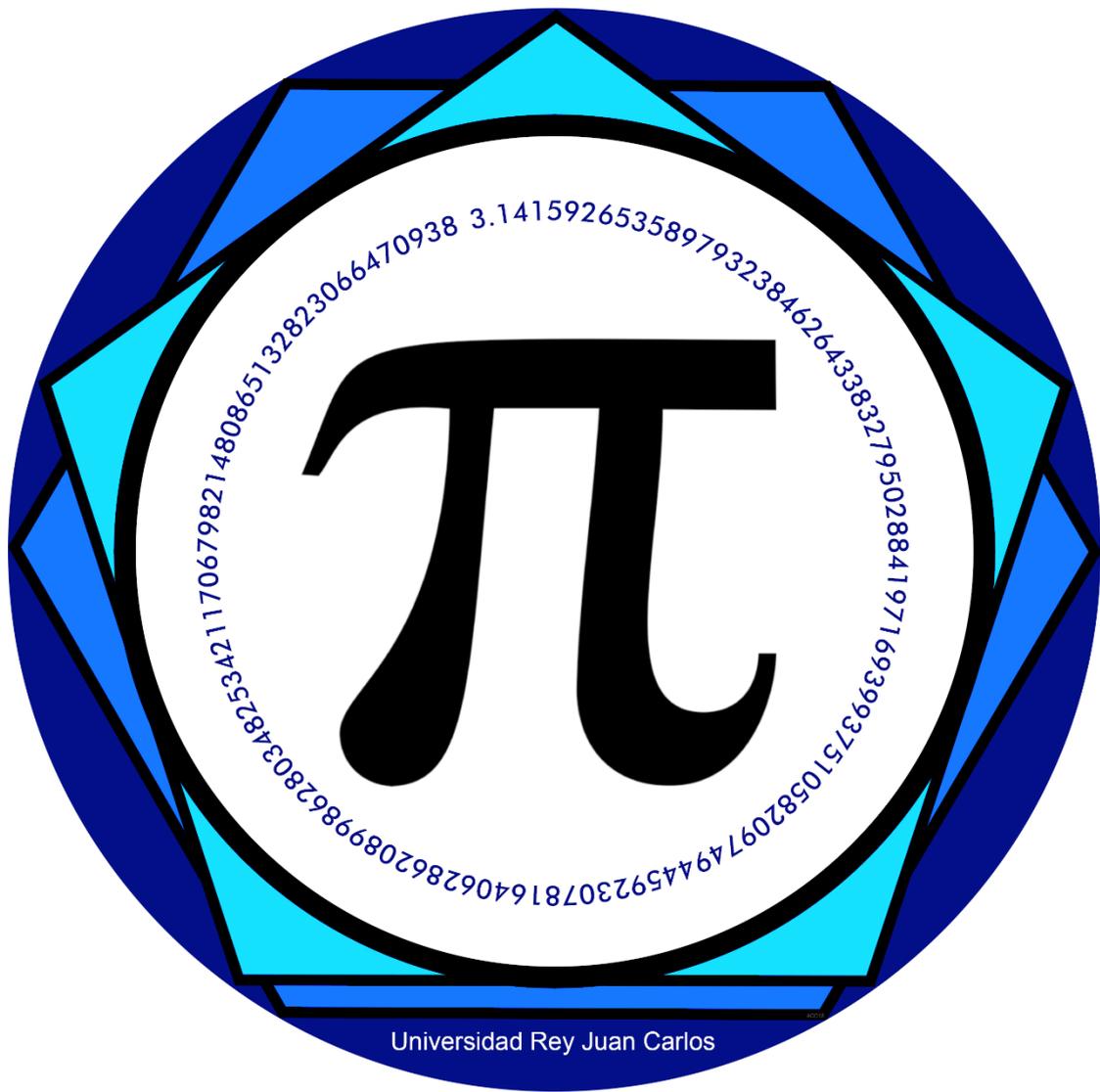


Ciencias Experimentales

# ACTAS DÍA DE PI

ISBN: 978-84-09-26470-4

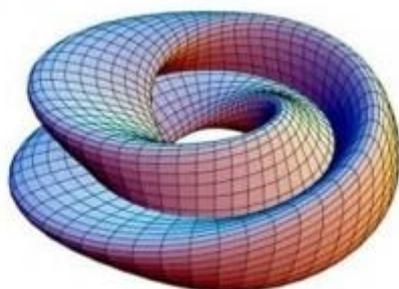


# Actas Día de $\pi$

## 2021

# Actas

---



**DIA DE  $\pi$  2021**

**12 de marzo de 2021**

**Universidad Rey Juan Carlos**

## El horario de la jornada ha sido el siguiente:

**9:00** Inauguración: *Abraham Duarte Muñoz, Esther García, Clara Simón de Blas* (URJC)

**9:30** Optimización Matemática para la mejora de la gestión industrial. *Antonio Alonso* (URJC)

**10:00** Análisis Bayesiano de Imágenes Médicas. *Raquel Montes* (URJC)

**10:30** Cuantificación de la incertidumbre, también sobre los modelos. *María Eugenia Castellanos* (URJC)

**11:00** Coffee break.

**11:30** Una mirada a las álgebras no asociativas. *Esther García* (URJC)

**12:00** Un mundo sin la RSME. *David Martín de Diego* (RSME)

**12:30** Análisis de redes complejas: Del cerebro a Google. *Miguel Romance* (URJC)

**13:00** Más humanos que los humanos. Sistemas discriminativos que no discriminan. *Alfredo Cuesta* (URJC)

**13:30** Modelización de sistemas dinámicos y evolutivos. *Michael Christian Stich* (URJC)

**14:00** Una introducción a las ecuaciones de difusión no lineal. *Razvan Gabriel Iagar* (URJC)

**14:30** Lunch break

**16:00** Taller: Lógica y programación funcional: dos caras de la misma moneda. *Juan Manuel Serrano* (URJC)

**17:00** Taller: Desafíos matemáticos. *Adrián Bacelo Polo* (URJC)

**18:00** Clausura

# Abstracts

**Día de  $\pi$  2021**

12 de marzo de 2021

Universidad Rey Juan Carlos, Móstoles

# Optimización Matemática para la mejora de la gestión industrial

Antonio Alonso-Ayuso, María Sierra Paradinas

Universidad Rey Juan Carlos

[antonio.alonso@urjc.es](mailto:antonio.alonso@urjc.es), [msierra@idom.es](mailto:msierra@idom.es)

## Resumen

Desde que en 1947 George Dantzig publicara su algoritmo del Simplex<sup>[1]</sup>, la optimización matemática se ha convertido en una herramienta fundamental en la gestión óptima de recursos. El desarrollo de la informática y las mejoras en los métodos de resolución han hecho posible que hoy en día sea posible resolver de modelos con millones de variables (muchas enteras) y restricciones en tiempos razonables.

El DSLAB, grupo de investigación de alto rendimiento de la URJC, lleva años colaborando con empresas de diversos sectores en la aplicación de esta disciplina para la mejora de sus procesos.. En esta charla introduciremos dos de las últimas colaboraciones:

- Planificación de la red de distribución de Repsol en Perú: esta compañía dispone de una refinería en Perú que suministra más del 50% de los productos petrolíferos del país. Su distribución se hace diariamente por barco de a uno de los seis centros de distribución costeros o por tren a los dos centros situados en el interior. Además, se sirve de miles de camiones que, además de ayudar en esta distribución, hacen

---

<sup>[1]</sup>Seleccionado como uno de los diez algoritmos más decisivos del siglo XX, “The top ten algorithms of the century”, suplemento a Computing in Science and Engineering, 1, 6, IEEE, 2000.

el reparto capilar a los clientes finales (entre los que se incluyen las más de 800 gasolineras de Repsol en Perú).

El DSLAB ha desarrollado dos modelos de optimización. El primero ayuda a la planificación estratégica: decidir cuántos barcos tiene disponibles cada mes del año y qué necesidades de importación van a necesitar Repsol para cubrir la demanda esperada. El resultado de este primer modelo se utiliza como entrada en el segundo modelo, que con un horizonte de planificación de varias semanas, determina el recorrido que debe hacer y la carga que debe llevar cada medio de transporte para satisfacer la demanda y minimizar los costes logísticos.

- Optimización de la operación de corte: en colaboración con la empresa Cortichapa, hemos desarrollado un modelo de optimización que permite decidir cómo deben cortarse las bobinas de acero con las que trabaja la empresa con el fin de poder atender la demanda solicitada, pero reduciendo el sobrante no utilizado de las bobinas. El modelo permite aprovechar el 80% del material utilizado, cuando la operativa actual reduce este porcentaje a algo menos del 60%.

**Día de  $\pi$  2021**

12 de marzo de 2021

Universidad Rey Juan Carlos, Móstoles

## **“Más humanos que los humanos”. Sistemas discriminativos que no discriminan**

Alfredo Cuesta Infante

E.T.S. Ing. Informática, URJC

alfredo.cuesta@urjc.es

### **Resumen**

En la actualidad, la cantidad de datos disponible, el incremento y abaratamiento de la potencia de cálculo y la irrupción de las técnicas de aprendizaje profundo han posibilitado una nueva generación de algoritmos con un rendimiento superior al humano en muchas tareas donde antes era casi impensable, especialmente en el campo del aprendizaje automático; es decir la rama de la inteligencia artificial que tiene como objetivo crear algoritmos a partir de conjuntos de datos o experiencias.

En este tipo de problemas se suele decir que la máquina *aprende*; y dados los buenos resultados es natural imaginar que la posibilidad de máquinas realmente inteligentes está cada vez más cerca.

Pero la realidad es muy distinta y menos fantasiosa. Una gran parte de los problemas de aprendizaje automático se reducen a construir una función matemática que sea capaz de discriminar la entrada entre varias opciones o clases posibles. Estas máquinas no son más que funciones con miles de parámetros, que no hacen otra cosa que interpolar a partir de los datos dados minimizando una cierta pérdida.

Estos datos, en muchas ocasiones, son un reflejo de la sociedad porque están extraídos directamente de ella. Si creamos algoritmos que afectan a la vida de la gente como la concesión de un crédito, la admisión en un

programa, la probabilidad de reincidencia, y los entrenamos con estos datos, es muy probable que los resultados hereden los sesgos que llevan los datos.

La comunidad de aprendizaje automático ha tomado conciencia de este problema recientemente y se ha realizado un gran esfuerzo para modelar la equidad (*fairness*) y así incorporarla al problema de minimización. Muy brevemente, la equidad se puede modelar a dos niveles: individual y grupal. La equidad individual se resume en el lema “*individuos similares deben recibir tratos similares*”. Por otro lado, la equidad grupal pretende imponer que “*todos los grupos reciban un trato similar*”. Por contraintuitivo que pueda parecer se puede demostrar que ninguna de las dos es completamente justa. En esta charla se expondrán estos y otros conceptos básicos, así como problemas actuales abiertos sobre esta línea de investigación. También se hace una breve demostración con la herramienta de Google research disponible en

<https://research.google.com/bigpicture/attacking-discrimination-in-ml/>.

## Referencias:

1. Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Rein-gold, and Richard Zemel. “*Fairness through awareness*”. In Proc. of the 3rd Innovations in Theoretical Computer Science Conf., ITCS '12, pp. 214–226, 2012.
2. Moritz Hardt, Eric Price, and Nathan Srebro. “*Equality of opportunity in supervised learning*”. In NeurIPS, 29, pp.3315–3323, 2016.
3. Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez-Rodriguez, and Krishna P. Gummadi. “*Fairness beyond dis-parate treatment and dis-parate impact: Learning classifica-tion without disparate mistreatment*. In Proc. of the 26th Int. Conf. on World Wide Web, pp. 1171–1180, 2017.

**Día de  $\pi$  2021**

12 de marzo de 2021

Universidad Rey Juan Carlos, Móstoles

## Una mirada a las álgebras no asociativas

Esther García

Departamento MACIMTE, URJC

esther.garcia@urjc.es

### Resumen

En esta charla haremos una breve incursión en el campo de las álgebras no asociativas. Nos apoyaremos en los conjuntos de matrices simétricas y matrices antisimétricas sobre un cuerpo para definir dos productos no asociativos, el producto Jordan y el producto Lie. Una vez definidos estos dos productos no asociativos, remarcaremos las principales identidades que se satisfacen, y a partir de aquí abstraeremos las definiciones de álgebra de Jordan y álgebra de Lie. Aunque hay muchos más ejemplos de estructuras no asociativas (octoniones, álgebras alternativas, Poisson algebras,...) en esta charla nos centraremos en estos dos modelos de álgebras no asociativas.

Haremos una introducción histórica de cada una de estas dos estructuras, mencionando a destacados matemáticos que han trabajado con ellas (algunos de ellos, premiados con una medalla Fields o con un premio Abel), y comentaremos la situación actual.

Aunque sus orígenes fueron independientes, hablaremos de distintas construcciones que conectan ambas estructuras, destacando el cuadrado mágico de Tits, que permitió dar ejemplos explícitos de algunos modelos de álgebras de Lie utilizando como ingredientes a las álgebras de Jordan, así como la construcción de Tits-Kantor-Koecher que construye un álgebra de Lie con una 3-graduación a partir de cualquier álgebra de Jordan.

Para concretar algo más mi trabajo y mis aportaciones en estos dos campos, hablaré de dos generalizaciones de la construcción de

Tits-Kantor-Koecher que permiten asociar estructuras de Jordan a álgebras de Lie que tienen elementos o subconjuntos destacados. El principal objetivo de introducir estas conexiones entre los “mundos” Lie y Jordan es intentar aprovechar resultados del ámbito Jordan para conocer mejor las álgebras de Lie, especialmente en el caso de dimensión infinita donde todavía queda mucho por hacer.

### Referencias:

1. A. Fernández López. *Jordan structures in Lie algebras*, vol. 240 in *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2019.
2. K. McCrimmon. *A taste of Jordan algebras*. Universitext. Springer-Verlag, New York, 2004.
3. N. Jacobson. *Lie algebras*, vol. 10 in *Interscience Tracts in Pure and Applied Mathematics*. Interscience Publishers, New York-London 1962.
4. N. Jacobson. *Lectures on quadratic Jordan algebras*, vol. 45 in *Tata Institute of Fundamental Research Lectures on Mathematics*. Tata Institute of Fundamental Research, Bombay, 1969.
5. H. Strade. *Simple Lie algebras over fields of positive characteristic (I) and (II)*, vols. 38 and 42 in *De Gruyter Expositions in Mathematics*. De Gruyter, Berlin, 2017.

**Día de  $\pi$  2021**

12 de marzo de 2021

Universidad Rey Juan Carlos, Móstoles

# Una introducción a las ecuaciones de difusión no lineal

Razvan Gabriel Iagar

Universidad Rey Juan Carlos, Madrid

razvan.iagar@urjc.es

## Resumen

Las ecuaciones de difusión no lineal aparecen de forma natural en numerosos modelos de la naturaleza. En esta conferencia se introducen algunos modelos clásicos de ecuaciones de difusión no lineal, en forma de variaciones no lineales de la ecuación del calor, presentando su motivación y modelización: la ecuación de los medios porosos y de difusión rápida

$$\partial_t u = \Delta u^m, \quad m > 1 \text{ medios porosos, } m \in (0, 1) \text{ difusión rápida,}$$

y respectivamente la ecuación parabólica p-Laplaciana

$$\partial_t u = \Delta_p u, \quad \Delta_p u(x, t) = \operatorname{div}(|\nabla u|^{p-2} \nabla u)(x, t), \quad p > 1.$$

Hablaré sobre los problemas más interesantes en relación a esta clase de ecuaciones, en particular aquellos relacionados con la dinámica de las ecuaciones y los modelos de comportamiento a largo plazo de sus soluciones. Se presentarán también muy brevemente los efectos de reacción o absorción y su influencia sobre la dinámica de las ecuaciones incluyendo términos de este tipo.

**Día de  $\pi$  2021**

12 de marzo de 2021

Universidad Rey Juan Carlos, Móstoles

## Lógica y programación funcional - dos caras de la misma moneda

Juan Manuel Serrano

Universidad Rey Juan Carlos, Madrid

juanmanuel.serrano@urjc.es

### Resumen

¿Qué tiene que ver la lógica con la computación? ¿Qué tienen que ver las conectivas proposicionales y las demostraciones lógicas, con los tipos de datos y los programas? Aparentemente, se trata de mundos completamente distintos, y de hecho fueron descubiertos y desarrollados de forma independiente. Pero la correspondencia de Curry-Howard nos enseña que la lógica proposicional intuicionista y el cálculo lambda que subyace a los lenguajes de programación funcional, son esencialmente la misma teoría: que cuando programamos estamos haciendo demostraciones, y que cuando estamos demostrando teoremas, programamos. Ilustraremos este hecho mostrando cómo utilizar el lenguaje de programación Scala para formalizar y demostrar acertijos lógicos.

### Referencias:

1. Juan Manuel Serrano. Slides y código. <https://github.com/jserrano/ohidalgo/1ahoradelambda>
2. Philip Wadler, Propositions as types. Commun. ACM 58(12): 75-84 (2015)

**Día de  $\pi$  2021**

12 de marzo de 2021

Universidad Rey Juan Carlos, Móstoles

## Los sonidos de $\pi$

Clara Simón de Blas<sup>1</sup> y Julio Flores<sup>2</sup>

URJC. Departamento de Ciencias de la Computación, Arquitectura de Computadores, Lenguajes y Sistemas Informáticos y Estadística e Investigación Operativa. Matemática Aplicada, Ciencia e Ingeniería de los Materiales y Tecnología Electrónica

clara.simon@urjc.es; julio.flores@urjc.es

### Resumen

En este trabajo se presenta y analiza la composición ideada por David Macdonald sobre el número  $\pi$ . Este músico, ideó una manera de asociar números a las teclas de un piano para interpretar melodías basadas en secuencias numéricas. El piano que creó Bartolomeo Cristofori contaba con 54 teclas. Según la música de piano fue evolucionando, el rango del teclado fue creciendo para satisfacer las necesidades de los compositores que buscaban ampliar su potencial expresivo. En torno a 1890, el rango actual de 88 teclas era ya un estándar, con  $7 \frac{1}{4}$  (de A0 a C8; 27,5 Hz a 4.186Hz\*). ¿Porqué no se aumentó la longitud del teclado? El oído humano puede detectar sonidos entre 20Hz y 20.000Hz, aunque el límite de frecuencias que el cerebro humano puede identificar está en el mejor de los casos sobre 4.000Hz. Incluso si considerásemos ampliar el rango del teclado (como hizo la mítica Bösendorfer con el modelo Imperial que cuenta con 97 teclas, 8 octavas completas y casi 3 metros de longitud), las teclas graves adicionales serían poco más que un murmullo ininteligible, mientras que las teclas agudas serían percibidas como disonancias desagradables. El método ideado por David Macdonald se basa en teclados de 88 teclas. Una vez creada la melodía principal, el acompañamiento se construye basándose en la teoría de armonía musical, utilizando la escala armónica de La menor.

Se finaliza este trabajo presentando algunas curiosidades sobre el número PI a lo largo de la historia y presentando una interpretación por los músicos Nicolás Flores y Julio Alberto Flores de la composición ideada por David Macdonald sobre el número PI.

### Referencias:

1. <https://www.hooktheory.com/theorytab/view/david-macdonald/song-from-pi?node=1.364.6>
2. <https://rolloid.net/descubre-como-suena-el-numero-pi/>

**Día de  $\pi$  2021**

12 de marzo de 2021

Universidad Rey Juan Carlos, Móstoles

# Modelización de sistemas dinámicos y evolutivos

Michael Stich

Departamento MACIMTE, URJC

michael.stich@urjc.es

## Resumen

Las propiedades fundamentales de la vida, las reacciones químicas, y los procesos físicos que organizan nuestro mundo dependen del paso del tiempo. Por tanto, la modelización matemática de estos sistemas dinámicos y evolutivos es un campo de interés fundamental para entender nuestro mundo y cambiarlo a través de las ingenierías.

Uno de los métodos más utilizados en este contexto son las ecuaciones diferenciales, tanto ecuaciones diferenciales ordinarias como en derivadas parciales. En esta contribución, revisamos varios ejemplos de relevancia y ponemos el énfasis en una aplicación en el área entre física, química y las ingenierías: la de entender, controlar y optimizar una reacción química, en concreto la oxidación de CO en el catalizador de platino (p.ej. utilizado en los coches de gasolina). La modelización matemática nos lleva a un modelo de tres ecuaciones en derivadas parciales no lineales que describe cualitativamente, y en algunos rangos de parámetros incluso cuantitativamente, las propiedades fundamentales del experimento, entre ellas la dinámica periódica de la reacción química. Este modelo se puede modificar según el planteamiento del método de control. Si, por ejemplo, se aplica una subida local de temperatura (p.ej., con un láser), la frecuencia de las oscilaciones cambia locamente y se observa la aparición de ondas concéntricas. En otro rango de parámetros se observa un acoplamiento global a través de la fase gaseosa y el sistema se homogeneiza, tanto en el experimento como en las simulaciones.

# Presentaciones



## **Cuantificación de la incertidumbre, también sobre los modelos**

**María Eugenia Castellanos**  
[maria.castellanos@urjc.es](mailto:maria.castellanos@urjc.es)

Día de Pi. Móstoles. 12 marzo 2021



Varios ejemplos motivadores

Procedimiento de selección de modelos (selección de variables)

Verosimilitud: ¿cómo se construye con datos censurados?

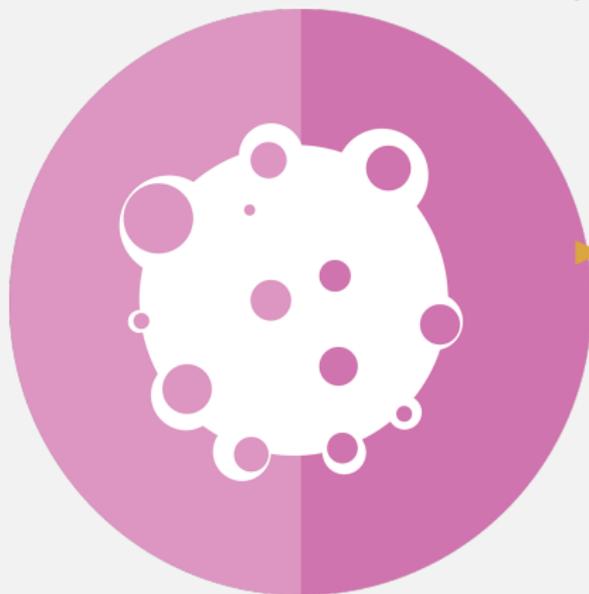
Cuestiones abiertas en los problemas con censura

Resultados en el ejemplo real

# Estudios de asociación de genoma con ciertas enfermedades (GWAS)



- ▶ Número  $p$  de genes, mucho mayor a  $n$ , número de individuos.
- ▶ Objetivo: qué genes aparecen diferentemente expresados por grupos: sanos/enfermos; o asociados a una variable continua.
- ▶ Aplicación a un desorden, la Beta-thalassemia, en una muestra bastante aislada de Talana, en Cerdeña (Italia).
- ▶ La variable respuesta en este caso es  $MCV = \text{media}$



- ▶  $n = 2116$  mujeres 2004-2013.
- ▶  $t_i$  = tiempo hasta el fallecimiento (en años) o censurado final estudio, **afortunadamente la mayoría!!**
- ▶ Relación de estas covariables con la supervivencia:
  - ▶ Nº nodos afectados;
  - ▶ Edad;
  - ▶ Recurrencia;
  - ▶ Metástasis;
  - ▶ Receptor hormonal de Estrógeno;
  - ▶ Receptor hormonal de Progesterona;



***“All models are wrong but some of them are useful”***

***George E.P. Box (1919-2013)***



$$Y = f(X, \theta) + \epsilon$$

**PROBLEMA DE SELECCIÓN DE VARIABLES:** disponemos de un conjunto inicial de potenciales variables explicativas

$$\mathbf{X} = (X_1, X_2, \dots, X_k)$$

**OBJETIVO:** Seleccionar las variables (genes) más relevantes para explicar la variabilidad en la respuesta,  $Y$ .

**MODELOS DE REGRESIÓN:**

$$Y = f(\mathbf{X}, \theta) + \epsilon = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon$$



Muchos nombres:

- ▶ Feature selection.
- ▶ Métodos de regularización.
- ▶ Selección de variables.
- ▶ Selección de modelos.

- ▶ Partir del modelo más complejo, y forzar que algunos coeficientes sean cero (LASSO, Ridge Regression).

$$f(\theta, \mathbf{X}) = \beta_0 + \beta_1 \mathbf{X}_1 + \beta_2 \mathbf{X}_2 + \beta_3 \mathbf{X}_3 + \epsilon$$

↓

$$f(\theta, \mathbf{X}) = \beta_0 + \mathbf{0} \mathbf{X}_1 + \beta_2 \mathbf{X}_2 + \beta_3 \mathbf{X}_3 + \epsilon$$

- ▶ Considerar todos los posibles modelos,  $2^k$ , y seleccionar el mejor(es) modelos.

- ▶ Partir del modelo más complejo, y forzar que algunos coeficientes sean cero (LASSO, Ridge Regression).

$$f(\theta, \mathbf{X}) = \beta_0 + \beta_1 \mathbf{X}_1 + \beta_2 \mathbf{X}_2 + \beta_3 \mathbf{X}_3 + \epsilon$$

- ▶ Considerar todos los posibles modelos,  $2^k$ , y seleccionar el mejor(es) modelos.

$k = 3$  variables

$$M_0 : Y = \beta_0 + \epsilon$$

$$M_1 : Y = \beta_0 + \beta_1 \mathbf{X}_1 + \epsilon$$

$$M_2 : Y = \beta_0 + \beta_2 \mathbf{X}_2 + \epsilon$$

$$M_3 : Y = \beta_0 + \beta_3 \mathbf{X}_3 + \epsilon$$

$$M_4 : Y = \beta_0 + \beta_1 \mathbf{X}_1 + \beta_2 \mathbf{X}_2 + \epsilon$$

$$M_5 : Y = \beta_0 + \beta_1 \mathbf{X}_1 + \beta_3 \mathbf{X}_3 + \epsilon$$

$$M_6 : Y = \beta_0 + \beta_2 \mathbf{X}_2 + \beta_3 \mathbf{X}_3 + \epsilon$$

$$M_7 : Y = \beta_0 + \beta_1 \mathbf{X}_1 + \beta_2 \mathbf{X}_2 + \beta_3 \mathbf{X}_3 + \epsilon$$

Problema de selección de variables → SELECCIÓN DE MODELOS.

- ▶ Ajustamos los  $2^k$  modelos de regresión.

$$M_1 : y_i = \beta_0 + \beta_1 X_{1,i} + \sigma \varepsilon_i, \quad \varepsilon_i \sim f(\cdot), \quad i = 1, \dots, n,$$

(podemos escribir todos los posibles modelos con variables indicadoras)

- ▶ Distintas densidades sobre  $\varepsilon_i \sim f(\cdot)$  nos permiten definir diferentes modelos paramétricos: normal, skew-normal, etc.;
- ▶ Usados en supervivencia: Weibull, Gamma, Lognormal, etc.

- ▶ Simplificamos, y consideramos por ahora sólo dos modelos:

$$M_0 : Y = \beta_0 + \epsilon$$

$$M_1 : Y = \beta_0 + \beta_1 X_1 + \epsilon$$

- ▶ Alternativamente

$$M_0 : \beta_1 = 0$$

$$M_1 : \beta_1 \neq 0$$



Desde un punto de vista Bayesiano podemos calcular las **probabilidades a posteriori** de cada modelo:

$$p(M_1 | \text{data}) = \frac{P(\text{data} | M_1)P(M_1)}{P(\text{data})}$$

- ▶ De cara a seleccionar variables: **USAMOS UN CRITERIO PARA SELECCIONAR MODELOS** (por ejemplo el modelo con mayor probabilidad a posteriori).
- ▶ Más interesante aún: **Realizar PROMEDIADO DE MODELOS**, usando las **PROBABILIDADES A POSTERIORI SOBRE CADA MODELO** sin quedarnos con un único modelo.

- ▶ No elegir un sólo modelo, realizar **el promedio de todos con respecto a la probabilidad a posteriori**.
- ▶ Numerosas ramas de la ciencia están incorporando esta forma de realizar predicciones (forecasting):
  - ▶ Dentro de la investigación médica, un grupo de trabajo en **incertidumbre sobre los modelos** dentro de la iniciativa STRATOS.
  - ▶ En ciencias sociales, mucha atención en este problema: En la revista *European Economic Review* en 2016 número especial titulado **Model uncertainty in economics**, actualmente en *Econometrics* número especial llamado **Bayesian and frequentist model averaging**.
  - ▶ Dentro de la simulación de modelos, por ejemplo en previsiones meteorológicas también se realiza promediado de modelos. (Li y otros 2017)

En nuestro caso,  $Y$  es el tiempo hasta que se experimenta el suceso y como hemos comentado está **SUJETO A CENSURA POR LA DERECHA** y es observada si y sólo si:

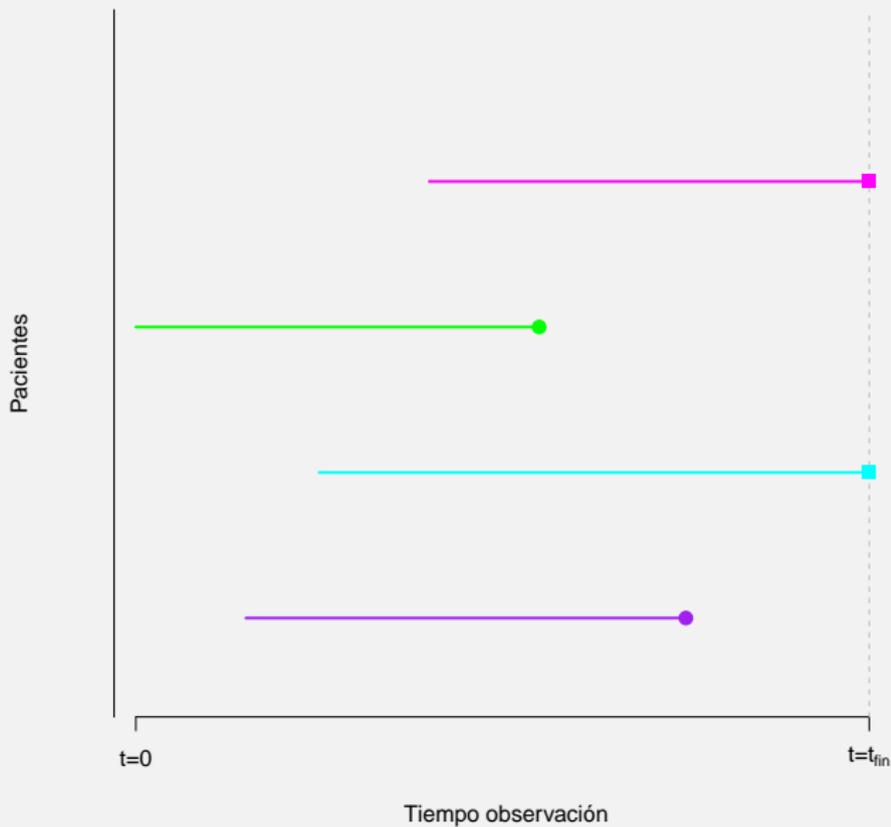
$$Y < c \quad (c \text{ tiempo en el que finaliza el estudio}).$$

De esta forma, en nuestro estudio, en el que tenemos  $n$  pacientes, la “información” contenida en cada unidad  $i \in \{1, 2, \dots, n\}$  varía:

Simplificando:

- ▶  $c_i$  pequeño  $\rightarrow y_i$  tiene menos opciones de ser observada, y por tanto el sujeto  $i$  contribuirá menos a la verosimilitud.
- ▶ Según  $c_i$  es mayor, observamos durante más tiempo a este paciente, el sujeto  $i$  contribuirá más a la verosimilitud.

# Tipos de pacientes y censura



Cada observación contribuye a la verosimilitud según:

- ▶ Para los sujetos donde se observa el tiempo,  $y_i$ :  $f(y_i)$ .
- ▶ En el caso de pacientes censurados, sabemos que  $y_i \geq c_i$ : la contribución a la verosimilitud es:

$$S(c_i) = 1 - F(c_i).$$

La verosimilitud se calcula como:

$$L = \prod_{i=1}^n f(y_i)^{\delta_i} S(c_i)^{1-\delta_i},$$

con  $\delta_i$  un indicador que toma los valores:

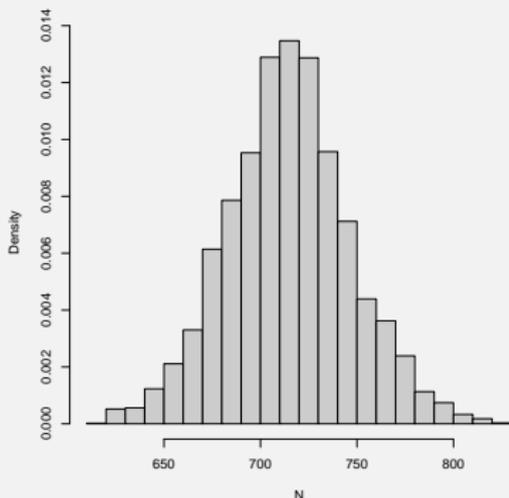
$$\delta_i = \begin{cases} 1 & \text{si } y_i \leq c_i, \text{ observamos el evento} \\ 0 & \text{si } y_i > c_i, \text{ observación censurada} \end{cases}$$



- ▶ **Construcción de la distribución inicial sobre los  $\beta$**  desde un punto de vista objetivo o “por defecto” para la selección de variables.
- ▶ Definición del tamaño *efectivo* de la muestra,  $n_{eff}$ . En los casos con censura,  $n$  no es el tamaño efectivo, tampoco lo es  $n_u...$
- ▶ **Aproximación numérica de las probabilidades a posteriori**, es necesario realizar integraciones en espacios de gran dimensión para calcular el denominador en la probabilidad a posteriori.
- ▶ **Proponer métodos de exploración del espacio de modelos mediante muestreo**, ya que no se pueden visitar todos si  $k$  es muy, muy grande.

Recordemos que teníamos:

- ▶  $n = 2116$  mujeres diagnosticadas,
- ▶  $y_i = \log(t_i)$ ,  $t_i =$  tiempo hasta el fallecimiento desde el diagnóstico, censurado muchas veces,
- ▶  $k = 6$  covariables: *número de nodos afectados, edad, recurrencia (0/1); metástasis (0/1); receptor hormonal de estrógeno (0/1); receptor hormonal de progesterona (0/1).*
- ▶  $n_u = 360$  completamente observadas, **83 % censura.**



- Promediado de la distribución a posteriori del tamaño efectivo de la muestra  $n_{eff}$  ( $E(n_{eff} | \text{data}) = 714$ ).
- Resumen:  $n_c = 1756$  'cuentan' como  $E(n_{eff} | \text{data}) - n_u = 354$  (20% de la información total)

# Resúmenes estándares que podemos conseguir usando selección de modelos

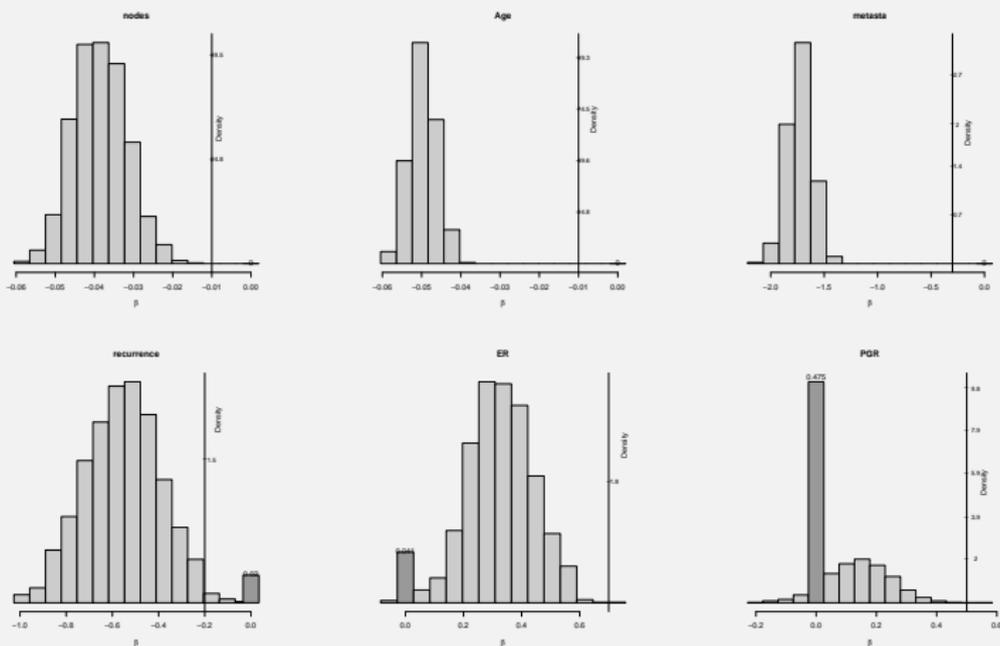


{nodos, edad, metastasis, recurrencia, ER, PGR}	0.473
{nodos, edad, metastasis, recurrencia, ER}	0.467

**Cuadro:** Probabilidad a posteriori para los dos modelos más probables.

Variable	nodes	age	metasta	recurrence	ER	PGR
Probability	1.00	1.00	1.00	0.98	0.96	0.52

**Cuadro:** Probabilidades de inclusión de cada variable



**Figura:** Distribución a posteriori para los coeficientes de regresión, promediada sobre todos los modelos.

# Predicciones usando promediado de modelos sobre la estimación de supervivencia



recurrencia	metastasis	nodos	edad	ER	PGR	Prob. superv. al	
						1	5
+	+	0	40	-	-	0.958	0.646
+	+	0	70	-	-	0.678	0.178
-	-	0	40	+	+	1	1
-	-	0	70	+	+	0.996	0.917
+	+	10	40	-	-	0.921	0.520
+	+	10	70	-	-	0.550	0.107
-	-	10	40	+	+	1	0.990
-	-	10	70	+	+	0.992	0.854
						0.999	0.941

**Cuadro:** Última fila es un caso promedio (valores de las covariables en la media)

- ▶ Problemas con **datos faltantes en las covariables**. Estamos empezando, no hay prácticamente nada hecho.
- ▶ Modelos construidos desde un punto de vista físico/biológico/matemático de simulación de una situación real, pueden haber muchos posibles modelos, ¿por qué quedarse sólo con uno? → **mucho mejor realizar promediado de modelos con probabilidades a posteriori**
- ▶ **Modelos de regresión con  $k$  gigante**,  $k \gg n$ , adaptar la metodología propuesta: búsqueda de modelos, definición distribuciones iniciales, etc.

# GRACIAS!!<sup>1</sup>

Más sobre mí y mi trabajo en:

<https://mecastellanos.wordpress.com/>

---

<sup>1</sup>This work has been supported by Ministerio de Ciencia e Innovación grant PID2019-104790GB-I00.

- ▶ M.J. Bayarri, J.O. Berger, A. Forte, and G. García-Donato. Criteria for Bayesian model choice with application to variable selection. *The Annals of Statistics*, 40:1550–1577, 2012.
- ▶ S. Cabras, M.E. Castellanos, and S. Perra. Comparison of objective Bayes factors for variable selection in parametric regression models for survival analysis. *Statistics in medicine*, 33(26):4637–4654, 2014.
- ▶ S. Cabras, M.E. Castellanos, and S. Perra. A new minimal training sample scheme for intrinsic bayes factors in censored data. *Computational Statistics & Data Analysis*, 81:52–63, 2015.
- ▶ M.E. Castellanos, G. García-Donato, S. Cabras. A Model selection approach for Variable selection with censored data, *Bayesian Analysis*, 16(1): 271-300 (2021).
- ▶ Li y otros 2017. *A Review on Statistical Postprocessing Methods for Hydrometeorological Ensemble Forecasting*

**Más** Sistemas  
**humanos** discriminativos  
**que los** que no  
**humanos** discriminen

2021.3.14

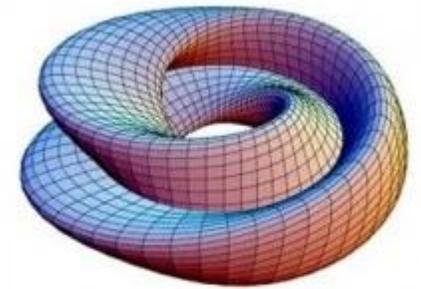
2021.3.12

Alfredo Cuesta



Universidad  
Rey Juan Carlos

i Escuela Técnica Superior de  
Ingeniería Informática



**DIA DE  $\pi$  2021**

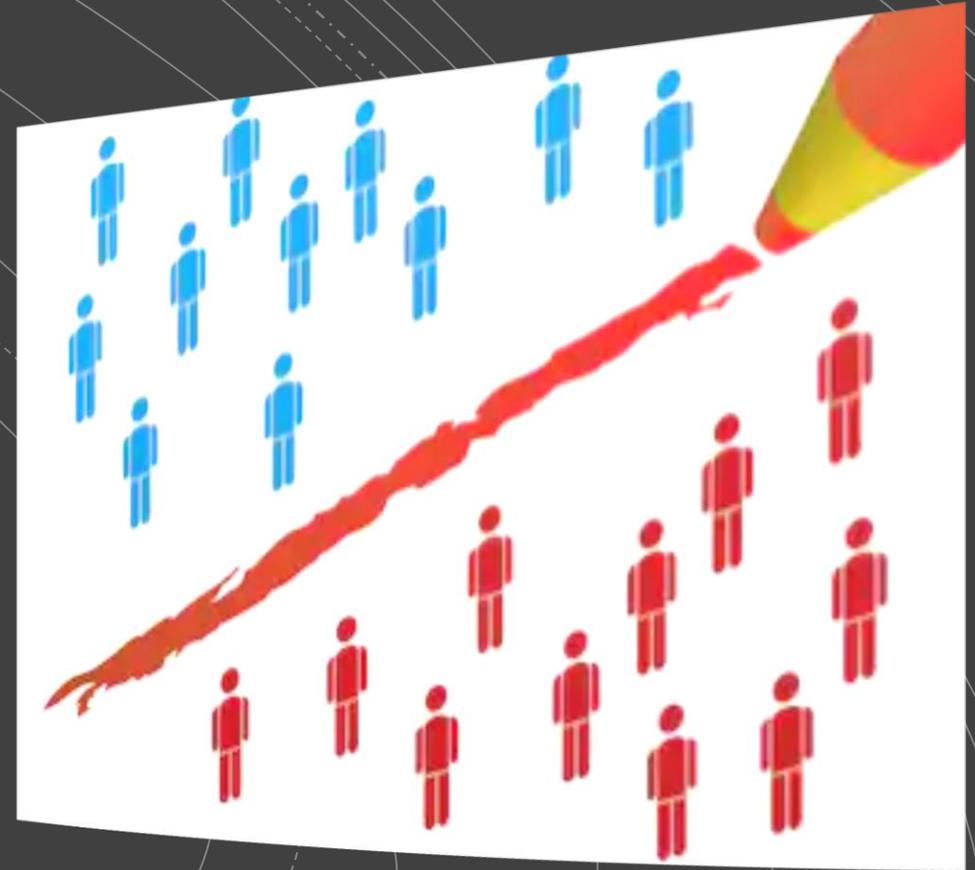


**Computación  
Avanzada  
Percepción y  
Optimización**

# Aprendizaje automático con sistemas discriminativos

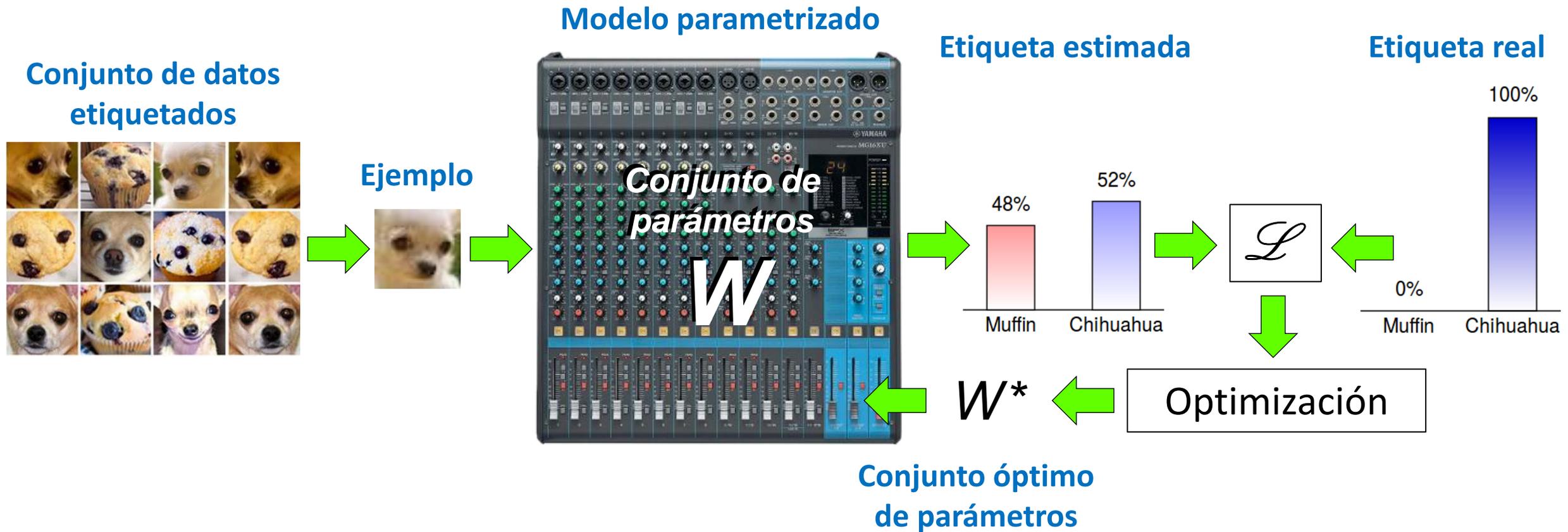


Así se vende



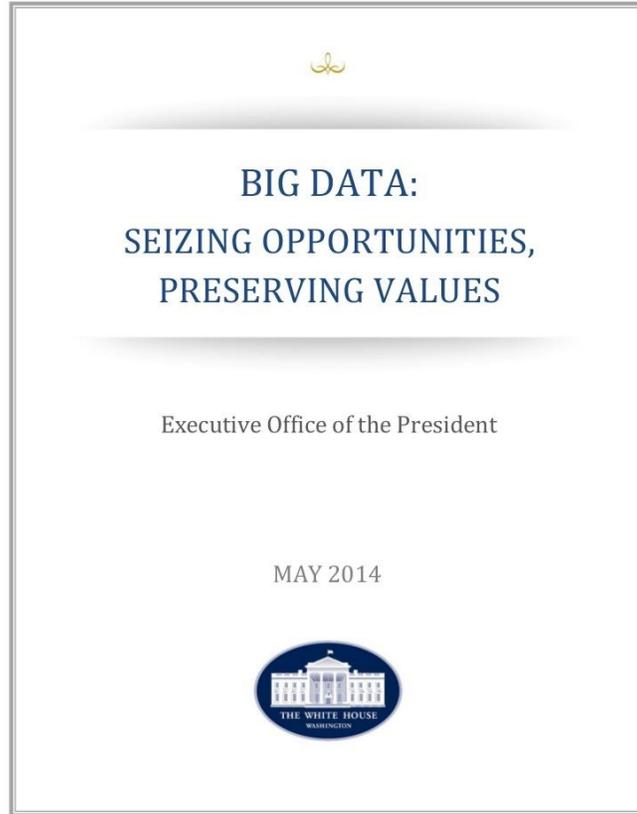
Pero sólo es un modelo paramétrico  
que optimiza una pérdida

# ¿Cómo se construye un sistema discriminativo?





Los datos  
son un reflejo  
de la sociedad



analyzed opportunities for technological innovation as well as primary challenges. One important social justice concern the 2014 report highlighted was “the potential of encoding discrimination in automated decisions”—that is, that discrimination may “be the inadvertent outcome of the way big data technologies are structured and used.” Building on these prior reports and the 2014 study conducted by the President’s Council of Advisors on Science and Technology (PCAST), the Administration is further examining how big data is used in the public

**BROOKINGS**

SERIES: A Blueprint for the Future of AI: 2018-2019

REPORT

**Credit denial in the age of AI**

Aaron Klein · Thursday, April 11, 2019

**MIT  
Technology  
Review**

Artificial Intelligence Dec 20, 2019

**A US government study confirms most face recognition systems are racist**

 **REUTERS**

BUSINESS NEWS OCTOBER 10, 2018 / 5:12 AM / A YEAR AGO

**Amazon scraps secret AI recruiting tool that showed bias against women**

Hay un algoritmo,  
que afecta a la vida de la gente,  
que está sesgado  
contra un grupo de población,



y tu  
pertenece  
a ese grupo



¿podemos diseñar sistemas discriminativos que no discriminen a grupos de población, a pesar de los sesgos en los datos?

**Más humanos  
que los humanos,  
ese es nuestro lema**

*Blade Runner, R. Scott 1982*

Si modelamos  
la **justicia**, la **equidad**,  
*(fairness)*  
podemos incorporarla  
al problema de  
optimización



## *Fairness* **individual**

0. *Through unawareness*

1. *Through awareness*



## *Fairness* **grupal**

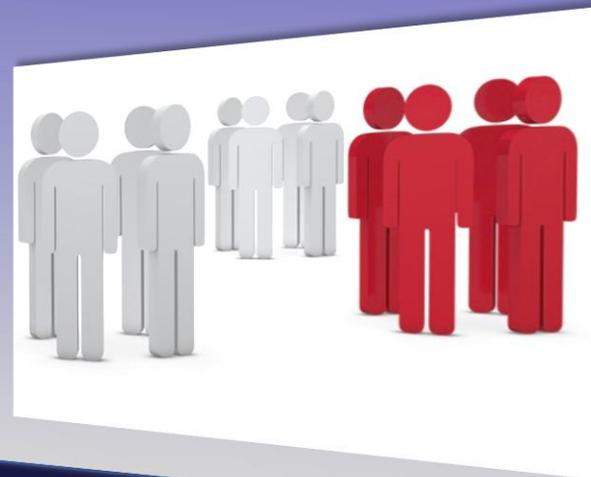
2. *Statistical parity*

3. *Predictive parity*

4. *Predictive equality*

5. *Equal opportunity*

6. *Equalized odds*



# Notación

**G** Atributo protegido o sensible

**X** Resto de atributos X

Cada individuo es un par (X,G)

**d** Asignación que genera el sistema discriminativo

**Y** “Etiqueta” = valor real asociado a (X,G)  
( también se le denomina *target* )

**G** define los grupos de población

- Grupo naranja
- Grupo azul



# 0. Fairness individual – *through unawareness*

Eliminar los atributos sensibles

- Intuitiva pero sin modelo matemático
- No evita *proxies*



The image shows a Europass Curriculum Vitae form for Davide Volpato. The form includes the Europass logo and a photo of the applicant, which is obscured by a large red 'X'. The personal information section contains the following details:

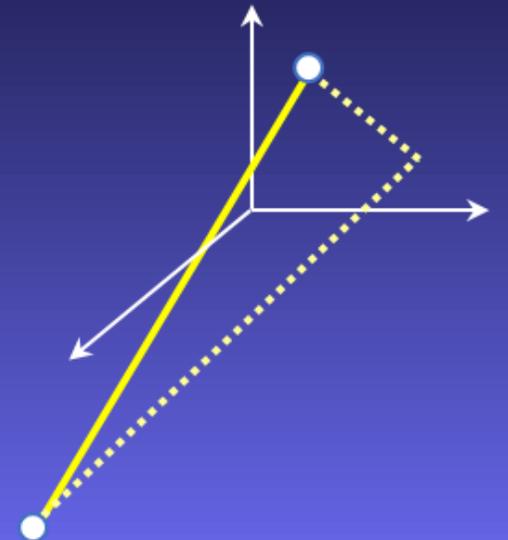
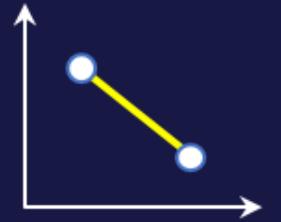
Personal information	<b>D.</b>
First name(s) / Surname(s)	<b>Da</b> <del>Volpato</del>
Address	Castello 3117, 30122 Venice (Italy)
Telephone(s)	(+39)041 713457
E-mail(s)	ilmondodavide@hotmail.com
Nationality	<del>It</del>
Date of birth	<del>1</del> <del>0</del> <del>87</del>
Gender	<del>M</del>

# 1. Fairness individual – *through awareness*

*“Individuos similares, tratos similares”*

→ necesitamos medir la similitud

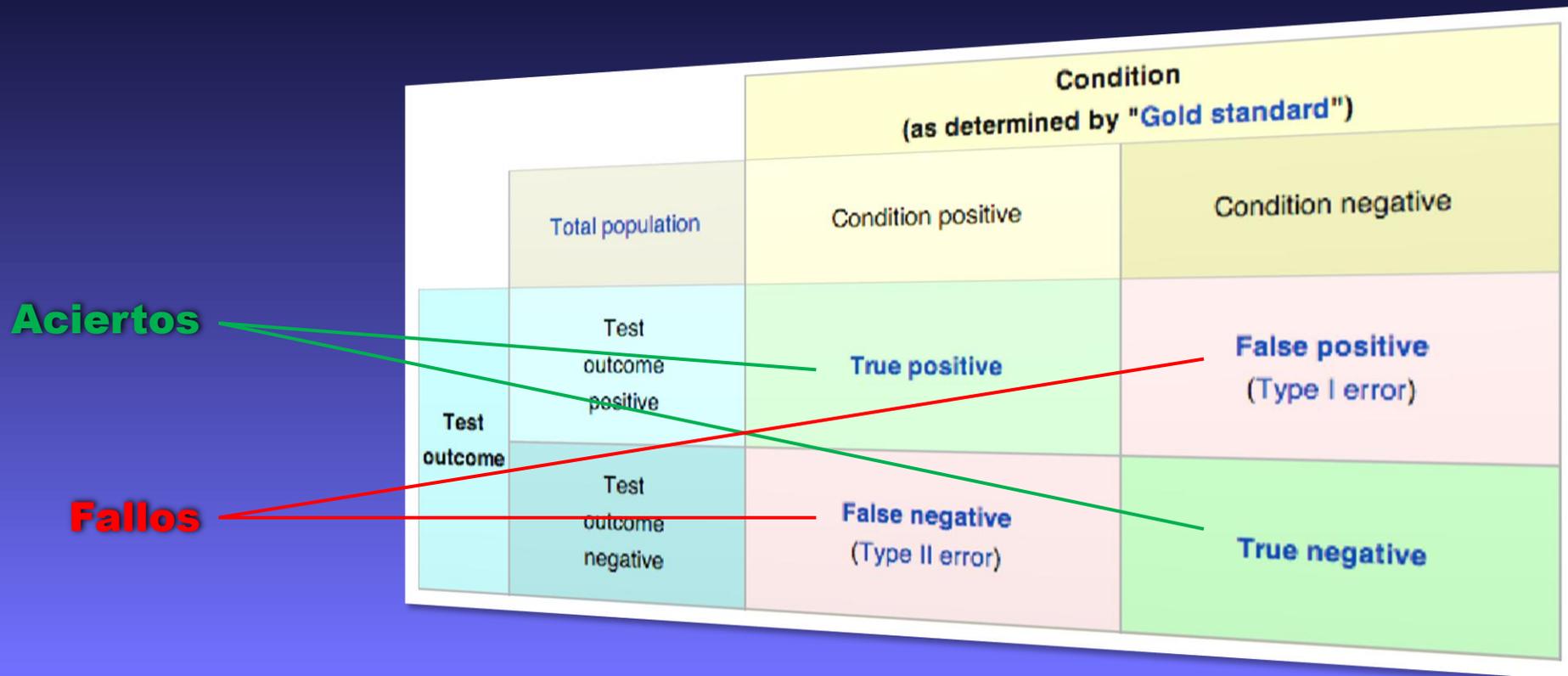
- Similitud  $\equiv$  Distancia
- La distancia depende de los atributos  
→ incurrimos en otro sesgo **más**  
según la información que recopilamos  
de cada individuo



# Fairness grupal

“Trato similar a cada grupo”

- Similitud  $\equiv$  valores estadísticos  $\approx$  **ratios** sobre los resultados
- Podemos utilizar los ratios derivados de la matriz de confusión



		Condition (as determined by "Gold standard")			
Total population		Condition positive	Condition negative	Prevalence = $\frac{\Sigma \text{ Condition positive}}{\Sigma \text{ Total population}}$	
Test outcome	Test outcome positive	True positive	False positive (Type I error)	Positive predictive value (PPV, Precision) = $\frac{\Sigma \text{ True positive}}{\Sigma \text{ Test outcome positive}}$	False discovery rate (FDR) = $\frac{\Sigma \text{ False positive}}{\Sigma \text{ Test outcome positive}}$
	Test outcome negative	False negative (Type II error)	True negative	False omission rate (FOR) = $\frac{\Sigma \text{ False negative}}{\Sigma \text{ Test outcome negative}}$	Negative predictive value (NPV) = $\frac{\Sigma \text{ True negative}}{\Sigma \text{ Test outcome negative}}$
Positive likelihood ratio (LR+) = TPR/FPR		True positive rate (TPR, Sensitivity, Recall) = $\frac{\Sigma \text{ True positive}}{\Sigma \text{ Condition positive}}$	False positive rate (FPR, Fall-out) = $\frac{\Sigma \text{ False positive}}{\Sigma \text{ Condition negative}}$	Accuracy (ACC) = $\frac{\Sigma \text{ True positive} + \Sigma \text{ True negative}}{\Sigma \text{ Total population}}$	
Negative likelihood ratio (LR-) = FNR/TNR		False negative rate (FNR) = $\frac{\Sigma \text{ False negative}}{\Sigma \text{ Condition positive}}$	True negative rate (TNR, Specificity, SPC) = $\frac{\Sigma \text{ True negative}}{\Sigma \text{ Condition negative}}$		
Diagnostic odds ratio (DOR) = LR+/LR-					

## 2. Fairness grupal – *statistical parity*

- Igualar la probabilidad de ser asignado a la clase positiva

$$\Pr (d=1 \mid G=a ) = \Pr (d=1 \mid G=n )$$

- **Variante:** *Conditional statistical parity*

Si existen factores **legítimos** (L) que condicionan la asignación positiva

$$\Pr (d=1 \mid L, G=a ) = \Pr (d=1 \mid L, G=n )$$

### 3. Fairness grupal – ***predictive parity***

- Igualar *Predictive Positive Values* (PPV)

$$\Pr ( Y=1 \mid d=1, G=a ) = \Pr ( Y=1 \mid d=1, G=n )$$

- Equivale a igualar *False Discovery Rates* (FDR)

$$\Pr ( Y=0 \mid d=1, G=a ) = \Pr ( Y=0 \mid d=1, G=n )$$

## 4. Fairness grupal – ***predictive equality***

- Igualar *False Positive Rates* (FPR)

$$\Pr ( d=1 | Y=0, G=a ) = \Pr ( d=1 | Y=0, G=n )$$

- Equivale a igualar *True Negative Rates* (TNR)

$$\Pr ( d=0 | Y=0, G=a ) = \Pr ( d=0 | Y=0, G=n )$$

## 5. Fairness grupal – **equal opportunity**

- Igualar *False Negative Rates* (FNR)

$$\Pr ( d=0 | Y=1, G=a ) = \Pr ( d=0 | Y=1, G=n )$$

- Equivale a igualar *True Positive Rates* (TPR)

$$\Pr ( d=1 | Y=0, G=a ) = \Pr ( d=1 | Y=0, G=n )$$

## 6. Fairness grupal – **equalized odds**

- Igualar *True Positive Rates* y *False Positive Rates* (TPR y FPR)

$$\Pr ( d=1 | Y=y, G=a ) = \Pr ( d=0 | Y=y, G=n )$$

para  $y = \{0,1\}$

# Google Research

*Attacking discrimination with  
smarter machine learning*

<http://>

Simulating loan decisions for different groups  
Drag the black threshold bars left or right to change the cut-offs for loans.  
Click on different preset loan strategies

Loan Strategy  
Maximize profit with:

MAX PROFIT

No constraints

GROUP UNAWARE

Blue and orange thresholds  
are the same

DEMOGRAPHIC  
PARITY

Same fractions blue / orange loans

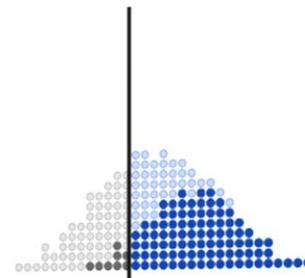
EQUAL  
OPPORTUNITY

Same fractions blue / orange loans  
to people who can pay them off

Blue Population

0 10 20 30 40 50 60 70 80 90 100

loan threshold: 50

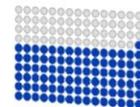


denied loan / would default granted loan / defaults  
denied loan / would pay back granted loan / pays back

Total profit = 19600

Correct 76%

loans granted to paying  
applicants and denied  
to defaulters

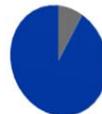


Incorrect 24%

loans denied to paying  
applicants and granted  
to defaulters



True Positive Rate 92%  
percentage of paying  
applications getting loans



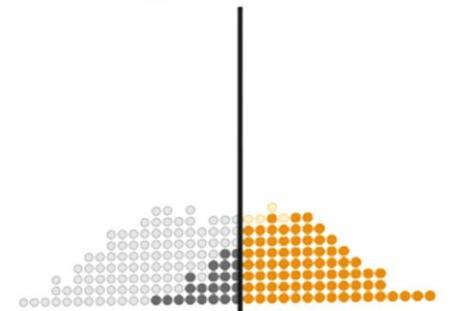
Positive Rate 66%  
percentage of all  
applications getting loans



Orange Population

0 10 20 30 40 50 60 70 80 90 100

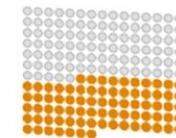
loan threshold: 50



denied loan / would default granted loan / defaults  
denied loan / would pay back granted loan / pays back

Correct 87%

loans granted to paying  
applicants and denied  
to defaulters

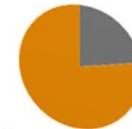


Incorrect 13%

loans denied to paying  
applicants and granted  
to defaulters



True Positive Rate 78%  
percentage of paying  
applications getting loans



Positive Rate 41%  
percentage of all  
applications getting loans



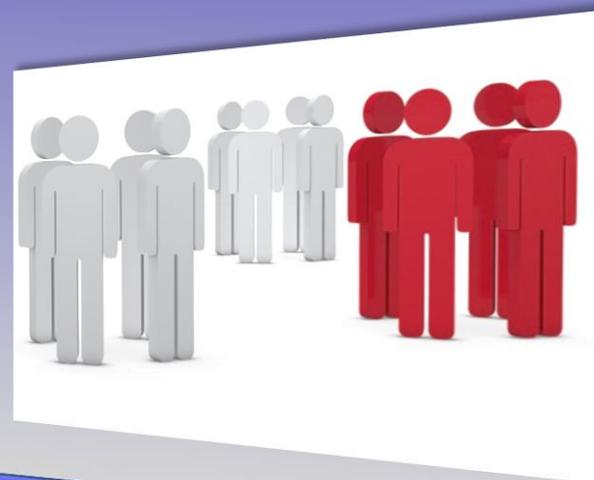
*Independencia*  $d \perp G$

*Suficiencia*  $Y \perp G \mid d$

*Separación*  $d \perp G \mid Y$

## Fairness **grupal**

2. *Statistical parity*
3. *Predictive parity*
4. *Predictive equality*
5. *Equal opportunity*
6. *Equalized odds*



¿Es posible cumplir las tres?

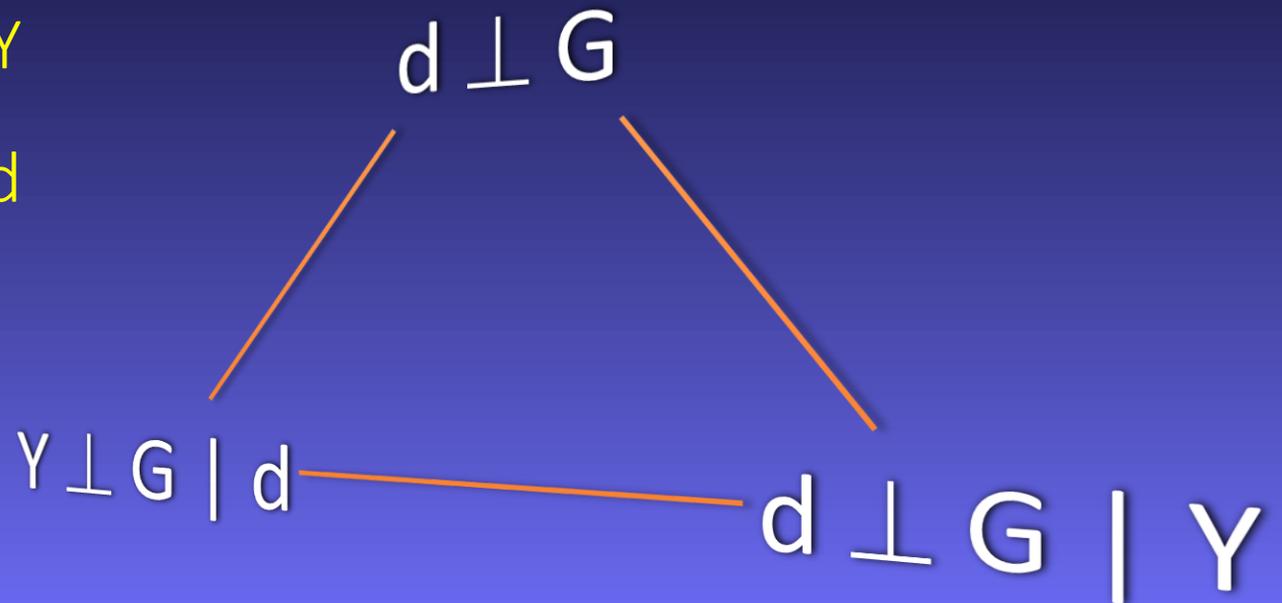
# Teorema de imposibilidad

Las tres definiciones son excluyentes dos a dos.

Si  $Y$  y  $G$  **no** son independientes, entonces o bien se logra

- $d \perp G$  o bien  $Y \perp G \mid d$
- $d \perp G$  o bien  $d \perp G \mid Y$
- $d \perp G \mid Y$  o bien  $Y \perp G \mid d$

pero no los dos a la vez



# Contribución

¿ Y si..

- Los clasificadores que vamos a utilizar no se pueden modificar,
- los individuos llegan uno tras otro
- y la etiqueta no está disponible hasta después de la predicción

?

*Esto ocurre en servicios en la nube que ofrecen MLaaS (Machine Learning as a Service)*

*“Towards Reducing Biases in Combining Multiple Experts Online”*

*Yi Sun , Iván Ramirez , Alfredo Cuesta , Kalyan Veeramachaneni*  
MIT , URJC

Enviado al IJCAI 2021

# Conclusiones

*Fairness* en algoritmos de ML es un problema **actual, abierto y necesario**

- Buscar un modelado “totalmente satisfactorio” → *Counterfactual fairness*
- Resolver problemas “reales”
  - Con datos perdidos
  - Con series temporales
  - De asignación de recursos
  - De aprendizaje no supervisado, por refuerzo...



**Más** Sistemas  
**humanos** discriminativos  
**que los** que no  
**humanos** discriminen



Universidad  
Rey Juan Carlos

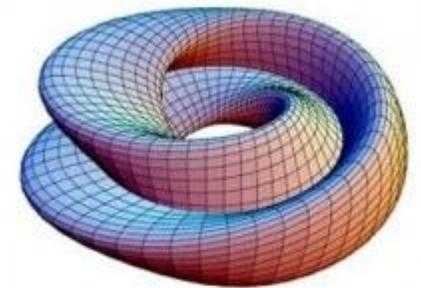
i Escuela Técnica Superior de  
Ingeniería Informática

Gracias por venir

2021.3.14

2021.3.12

[alfredo.cuesta@urjc.es](mailto:alfredo.cuesta@urjc.es)



**DIA DE  $\pi$  2021**



**Computación  
Avanzada  
Percepción y  
Optimización**

# Una mirada a las álgebras no asociativas

Esther García

Universidad Rey Juan Carlos

Móstoles, 12 de marzo de 2021

# ¿Qué es un álgebra?

Si  $F$  es un cuerpo, un *álgebra*  $A$  sobre  $F$  es un espacio vectorial con un “producto bilineal” (compatible con  $+$  y con  $\cdot_F$ )

$$A \times A \rightarrow A$$

- los elementos de  $A$  se pueden multiplicar entre sí

Ej: las matrices  $M_n(F)$ , donde  $F$  es un cuerpo:

$$A + B, \quad \lambda A, \quad A \cdot B$$

(el producto de matrices es **asociativo**:  $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ )

# Ejemplos de productos no asociativos

Si  $\frac{1}{2} \in F$  toda matriz  $A \in M_n(F)$  se puede escribir como

$$A = \frac{1}{2} \underbrace{(A + A^t)}_{\text{simétrica}} + \frac{1}{2} \underbrace{(A - A^t)}_{\text{antisimétrica}}$$

es decir

$$M_n(F) = \underbrace{\text{Sym}(M_n(F), t)}_H \oplus \underbrace{\text{Skew}(M_n(F), t)}_K$$

- $A, B \in H \not\Rightarrow A.B \in H$
- $A, B \in K \not\Rightarrow A.B \in K$

# Un producto bilineal en $H$

$$A, B \in H$$

$$A \bullet B = \frac{1}{2}(AB + BA)$$

- obs 1: si  $A = A^t$ ,  $B = B^t$  entonces  $(AB + BA)^t = AB + BA$
- obs 2: si  $A = A^t$  entonces  $A \bullet I = \frac{1}{2}(AI + IA) = A$  donde  $I$  es la matriz identidad
- obs 3: si  $A = A^t$ ,  $A^2 = A \bullet A$

$(H, \bullet)$  es un álgebra con elemento identidad

¡Ojo!: el producto  $\bullet$  no es asociativo:  $A \bullet (B \bullet C) \neq (A \bullet B) \bullet C$

# El producto no es asociativo

Contraejemplo:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

$$A \bullet (B \bullet C) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

$$(A \bullet B) \bullet C = \begin{pmatrix} 1 & \frac{3}{2} \\ \frac{3}{2} & 0 \end{pmatrix}$$

# $H$ con el producto $\bullet$

Propiedades del álgebra  $(H, \bullet)$ :

- No es asociativa
- Es conmutativa:  $A \bullet B = B \bullet A$
- Verifica la siguiente identidad (no muy intuitiva)

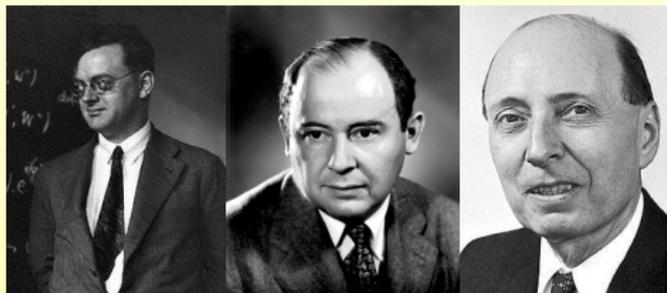
$$(A \bullet B) \bullet (A \bullet A) = A \bullet (B \bullet (A \bullet A))$$

(identidad de Jordan)

En realidad  $(H, \bullet)$  es un ejemplo de álgebra de Jordan

# Los “inventores”: P. Jordan, J. Von Neumann, E. Wigner

“On an algebraic generalization of the quantum mechanical formalism”. P. Jordan, J. V. Neumann, E. Wigner. *Annals of Mathematics* (1934)



# ¿Para qué se definen las álgebras de Jordan?

(1930') se busca estructura para expresar la Mecánica Cuántica

- los “observables” se representan mediante matrices simétricas
- ¿qué propiedades de los observables se pueden abstraer?

propiedad conmutativa + identidad de Jordan

Objetivo: encontrar objetos matemáticos (dim. infinita) nuevos

- Mejor si no están relacionados con productos asociativos
- ¿Habrán álgebras de Jordan que no vengan de álgebras asociativas?

# Desde el principio... esta es una historia de fracaso

En 1934 A. Albert encuentra un álgebra de Jordan *no asociativa*, pero demasiado pequeña (dimensión 27).

"On a Certain Algebra of Quantum Mechanics". A. Albert. *Annals of Mathematics*



# y... termina como una historia de fracaso (Física)

E. Zelmanov, Medalla Fields 1994.

- Clasifica las álgebras de Jordan (1980') y las usa para resolver un problema de teoría de grupos (Pr. Restringido de Burnside).
- Demuestra que el álgebra de Albert es la única “no asociativa”.



# En Matemáticas... las álgebras de Jordan sí tienen éxito

Desde su introducción (1930') llaman la atención de los algebraistas

- N. Jacobson: primer acercamiento infinito-dimensional
- K. McCrimmon: define las álgebras de Jordan cuadráticas
- E. Zelmanov: las clasifica en general
- J. Tits: las relaciona con las álgebras de Lie



# Otro producto no asociativo

Volvamos a las matrices  $M_n(F)$  y, en particular, a las antisimétricas

Un producto bilineal en  $K$

$A, B \in K$

$$[A, B] := AB - BA$$

- obs: si  $A = -A^t$ ,  $B = -B^t$  entonces  
 $(AB - BA)^t = -(AB - BA)$

$(K, [ , ]) es un álgebra$

**¡Ojo!**: el producto  $[ , ]$  no es asociativo



# $K$ con el producto $[ , ]$

Propiedades del álgebra  $(K, [ , ])$  :

- No es asociativa
- Es anti-conmutativa:  $[A, B] = -[B, A]$
- Verifica la siguiente identidad:

$$[A, [B, C]] = [[A, B], C] + [B, [A, C]]$$

(identidad de Jacobi)

En realidad  $(K, [ , ])$  es un ejemplo de álgebra de Lie

# Origen de las álgebras de Lie

- Sophus Lie (1842-1899): grupos de Lie  $\rightarrow$  álgebras de Lie
- Wilhelm Killing (1847-1923) (de modo independiente)  
clasificación (aprox.) de las álgebras de Lie de dim. finita sobre  $\mathbb{C}$



# Situación actual en álgebras de Jordan y de Lie

## Álgebras de Jordan

- Las álgebras de Jordan se conocen mejor que las Lie
- Clasificación completa en dimensión infinita

## Álgebras de Lie

- Clasificación completa en dimensión finita (reciente)
- Intentos de generalización a dim. infinita
- No se conoce la descripción de las simples en dim. infinita

## Referencias destacadas en Jordan y en Lie

- N. Jacobson. *Lie algebras*. Interscience Publishers, New York-London 1962.
- N. Jacobson. *Lectures on quadratic Jordan algebras*. Tata Institute of Fundamental Research, Bombay, 1969.
- K. McCrimmon. *A taste of Jordan algebras*. Universitext. Springer-Verlag, New York, 2004.
- H. Strade. *Simple Lie algebras over fields of positive characteristic (I) and (II)*. De Gruyter, Berlin, 2017.

# Cuadrado mágico de Freudenthal-Tits

Usando como “ingredientes” álgebras de Jordan se consiguen modelos de álgebras de Lie en dimensión finita. Destacamos  $E_7$  y  $E_8$ .

THE FREUDENTHAL–TITS MAGIC SQUARE:  $\mathcal{FT}(C, J)$

$C \setminus J$	$\mathbb{R}$	$\mathcal{H}_3(\mathbb{R})$	$\mathcal{H}_3(\mathbb{C})$	$\mathcal{H}_3(\mathbb{H})$	$\mathcal{H}_3(\mathbb{K})$
$\mathbb{R}$	0	$A_1$	$A_2$	$C_3$	$F_4$
$\mathbb{C}$	0	$A_2$	$A_2 \oplus A_2$	$A_5$	$E_6$
$\mathbb{H}$	$A_1$	$C_3$	$A_5$	$A_6$	$E_7$
$\mathbb{K}$	$G_2$	$F_4$	$E_6$	$E_7$	$E_8$

# Conexión Lie-Jordan: construcción TKK

Se debe a J. Tits, I. Kantor y M. Koecher (1962, 1964, 1967)



# La construcción TKK

Si  $J$  es un álgebra de Jordan, entonces

$$L = \underbrace{L_{-1}}_J \oplus \underbrace{L_0}_{[J,J]} \oplus \underbrace{L_1}_J$$

se convierte en un álgebra de Lie (3-graduada)  
Y recíprocamente...

$$L = L_{-1} \oplus L_0 \oplus L_1$$

los extremos  $(L_{-1}, L_1)$  tienen una estructura Jordan (par de Jordan)

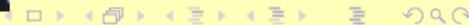
# ¿En qué temas trabajo yo?

Conexiones Lie-Jordan tipo TKK (en los extremos de un álgebra de Lie graduada hay una estructura Jordan)

Estructuras de Jordan asociadas a álgebras de Lie

- asociadas a elementos ad-nilpotentes
- asociadas a ideales internos abelianos

colaboradores: E. Neher, A. Fernández López, M. Gómez Lozano



# Otras estructuras no asociativas

- Álgebras alternativas (ej: octoniones)
- Álgebras “power-associative” (ej: sedeniones)
- Álgebras genéticas
- ...
- Generalizaciones del producto bilineal a otros “productos”:
  - Generalizaciones de álgebras de Jordan: pares y sistemas triples
  - Generalizaciones de Lie: triples de Lie
  - Pares de Kantor
  - Álgebras de Poisson
  - Álgebras de Jordan-Lie
  - ...

# Una introducción a las ecuaciones de difusión no lineal

**Razvan Gabriel Iagar**  
**Univ. Rey Juan Carlos, Madrid, Spain**

Día del Pi, Univ. Rey Juan Carlos, Madrid, 12 de marzo de 2021

# Outline

- 1 La ecuación de los medios porosos
- 2 Otros ejemplos notables de ecuaciones de difusión no lineal
- 3 Los principales problemas que se estudian

# La ecuación de los medios porosos

# La ecuación de los medios porosos

- Todos hemos estudiado la ecuación del calor

$$\partial_t u = \Delta u,$$

usando métodos diversos de resolución (separación de variable, transformada de Fourier etc.). Esta ecuación modeliza la difusión del calor (Fourier, 1822) y es el ejemplo fundamental de ecuación de difusión (lineal). Una solución de la misma es una función  $u(x, t)$  que satisface la ecuación en cierto sentido (clásico, fuerte, débil etc.) con  $x \in \mathbf{R}^N$  (o  $x \in \Omega \subset \mathbf{R}^N$ ) y  $t \geq 0$ .

- Entre las soluciones de la ecuación del calor, una muy destacada es el núcleo de Gauss

$$G(x, t) = \frac{1}{(4\pi t)^{N/2}} e^{-|x|^2/4t}, \quad x \in \mathbf{R}^N, t \geq 0,$$

y con la ayuda de esta solución, se puede resolver de forma semi-explícita la ecuación del calor mediante convolución.

- Observamos que esta solución tiene una forma particular muy importante, más precisamente

$$G(x, t) = t^{-\alpha} f(|x|t^{-\beta}), \quad \text{donde } \alpha = \frac{N}{2}, \beta = \frac{1}{2}.$$

Las soluciones de esta forma reciben el nombre de **soluciones auto-semejantes**.

# La ecuación de los medios porosos

- El primer ejemplo muy importante de ecuación de difusión no lineal es la **ecuación de los medios porosos** (abreviado PME)

$$\partial_t u = \Delta u^m,$$

con  $m > 1$ . Observamos que para  $m = 1$  recuperamos la ecuación del calor, mientras que para  $m < 1$  se conoce como la **ecuación de difusión rápida**.

- Modelización: describe el flujo de gas a través de un medio poroso (Leibenzon 1930, Muskat 1937). Este flujo se puede expresar en términos de la densidad en cada punto  $\varrho(x, t)$ , la presión  $p(x, t)$  y el campo vectorial de las velocidades  $\mathbf{V}(x, t)$ .
- La primera ley que se usa es la ley de continuidad de la mecánica de fluidos:

$$\epsilon \partial_t \varrho + \operatorname{div}(\varrho \mathbf{V}) = 0,$$

donde  $\epsilon \in (0, 1)$  es la porosidad del medio.

# La ecuación de los medios porosos

- La segunda ley fundamental usada en la teoría de la filtración de fluidos (y también muy útil en muchos otros fenómenos como modelos matemáticos del cáncer etc.) es la **ley de Darcy** (obtenida experimentalmente por primera vez por Darcy, 1856)

$$\mu \mathbf{V} = -k \nabla p,$$

donde  $k$  es la permeabilidad del medio,  $\mu$  es la viscosidad del fluido. Es la ecuación que sustituye las ecuaciones de Navier-Stokes para este tipo de medio.

- La ecuación de estado,  $p = p_0 \varrho^\gamma$ , donde  $\gamma \geq 1$  es el exponente de la transformación del gas ideal, isotérma si  $\gamma = 1$  o adiabática si  $\gamma > 1$ .
- Juntando todas estas ecuaciones, obtenemos la PME (con la incógnita  $u = \varrho$  la densidad del fluido), donde en este caso  $m = 1 + \gamma$  y en términos matemáticos simplificamos las constantes del medio poniéndolas iguales a 1.
- Otras aplicaciones muy conocidas: la propagación del calor por radiación en plasma (Zeldovich, Raizer, 1966), la filtración de un fluido a través de un medio poroso, típicamente la infiltración del agua en la tierra (problema de Boussinesq, 1903, y modelo de Polubarinova-Kochina, 1962), modelos de la dinámica de población etc.



# La ecuación de los medios porosos

- Observamos que las soluciones ZKB son también **auto- semejantes** y a la vez **con soporte compacto** en cualquier instante de tiempo. Difieren del núcleo de Gauss que tiene soporte infinito (cola decreciente cuando  $|x| \rightarrow \infty$ ).
- Eso refleja **el frente de avance** de la difusión (del gas, calor, agua infiltrada etc.): el soporte significa la zona con densidad positiva (ha llegado el frente) mientras que la zona donde vale 0 es la zona todavía "no infiltrada".
- En general, tratando de resolver la PME con un cierto dato inicial  $u_0$  (la densidad inicial del fluido cuya difusión se estudia), nos encontramos con un **problema de frontera libre**: la frontera de la zona ocupada por el fluido varía en cada instante del tiempo y su frontera no se conoce de antemano (es a su vez una incógnita del problema).
- Siendo una ecuación **no lineal**, todos los métodos lineales usados para resolver la ecuación del calor no se pueden usar para la PME y la ecuación no se puede resolver de forma explícita. Lo que se hace es estudiar las propiedades (funcionales, analíticas, geométricas, asintóticas) de las soluciones.



# La ecuación de difusión $p$ -Laplaciana

- Se trata de otra ecuación de difusión no lineal muy conocida y estudiada (abreviada PLE)

$$\partial_t u = \Delta_p u, \quad \Delta_p u(x, t) = \operatorname{div}(|\nabla u|^{p-2} \nabla u)(x, t),$$

el operador  $\Delta_p u$  se conoce como **el operador  $p$ -Laplaciano**. Aquí  $p > 1$  y para  $p = 2$  reconocemos la ecuación del calor.

- Aparece en modelos geométricos y en el estudio de los fluidos **non-Newtonianos**, es decir aquellos fluidos cuya viscosidad varía con la temperatura y la tensión (por ejemplo la sangre, plastilina, polimeros de todo tipo como resinas etc.)
- Es una ecuación de difusión no lineal que solo **depende del gradiente**: se escribe como

$$\partial_t u = \operatorname{div}(a(|\nabla u|) \nabla u), \quad a(|\nabla u|) = |\nabla u|^{p-2}.$$

- De nuevo, para  $p \in (1, 2)$  se conoce como **difusión rápida  $p$ -Laplaciana** y es una ecuación singular. Para  $p > 2$  es una ecuación degenerada, y el carácter degenerado o singular se observa esta vez en los puntos críticos de sus soluciones.

# La ecuación de difusión $p$ -Laplaciana

- Soluciones fundamentales de la ecuación de difusión  $p$ -Laplaciana, similares a las ZKB (se suelen llamar de la misma forma):

$$U(x, t) = t^{-\alpha} \left( D - k \left( \frac{|x|}{t^\beta} \right)^{p/(p-1)} \right)_+^{(p-1)/(p-2)},$$

donde  $\alpha = N\beta = N/(Np - 2N + p)$

- También son soluciones auto-semejantes y con soporte compacto. De la misma forma que para la PME, la ecuación  $p$ -Laplaciana da lugar a problemas de frontera libre.
- Se ha observado que muchas propiedades funcionales y ejemplos de soluciones notables de la PLE se parecen a las de la PME. Eso se ha podido justificar por la existencia de **una transformación** (cambio de variable) entre soluciones radialmente simétricas de la PME y de la PLE (R.I., A. Sánchez, J. L. Vázquez, 2008). Es una transformación que asocia  $p$  de la PLE con  $m + 1$  de la PME pero con cambio de dimensión espacial y de la variable independiente.
- Las propiedades funcionales (regularidad etc.) son más "ricas" en el caso de la PLE.

# Ecuaciones de reacción-difusión

- Se trata de ecuaciones "competitivas" entre un término de difusión y un término de reacción, teniendo como prototipo

$$\partial_t u = \Delta u^m + f(x, u),$$

donde lo más habitual es considerar  $f(x, u) = u^p$  (y más recientemente  $f(x, u) = |x|^\sigma u^p$ ) o  $f(u) = u(1 - u^p)$  (Fisher-KPP)

- La reacción **introduce "masa" en la ecuación**. Matemáticamente, la "masa" de la solución significa

$$M(t) = \int_{\mathbf{R}^N} u(x, t) dx.$$

- En las ecuaciones de difusión no lineal estándar (PME, PLE) la masa se conserva a lo largo de la evolución:  $M(t) = cte$ . Si tenemos reacción, la masa crece.
- Modelizan reacciones químicas (donde los reactantes no entran en la reacción de forma completa, sino que una parte se mueve en el medio de reacción debido a la difusión) y también dinámica de poblaciones (**ecuación logística** o Fisher-KPP para  $p = 1$ ).
- Las de tipo Fisher-KPP suelen tener soluciones en forma de **ondas viajeras**

$$u(x, t) = f(x - ct), \quad c = \text{velocidad de la onda.}$$

# Ecuaciones de absorción-difusión

- Se trata de ecuaciones "competitivas" entre un término de difusión y un término de absorción, teniendo como modelo

$$\partial_t u = \Delta u^m - f(x, u),$$

donde lo más habitual es considerar  $f(x, u) = u^q$ .

- En este caso la absorción **quita "masa" de la ecuación**: la masa total  $M(t)$  decrece con el tiempo.
- Un problema importante es si las soluciones llegan a "desaparecer": el decaimiento de la masa es tan fuerte que en algún momento, la solución con dato inicial  $u_0 \not\equiv 0$  se cancela: existe  $T > 0$  tal que  $u(T) \equiv 0$ . Esta situación se conoce como **extinción en tiempo finito**.
- Aparecen en muchos modelos de dinámica de poblaciones o en modelos matemáticos de evolución de tumores, y también en el estudio de reacciones químicas.

# Ecuaciones de absorción-difusión y reacción-difusión con gradientes

- Se trata de otra clase de ecuaciones de difusión no lineal "competitivas", donde el término de absorción o de reacción es de la forma

$$|\nabla u|^p.$$

- Se han estudiado mucho en conjunción con la ecuación  $p$ -Laplaciana debido a propiedades funcionales similares (siendo una ecuación que depende solo del gradiente):

$$\partial_t u = \Delta_p u + |\nabla u|^q, \text{ o } \partial_t u = \Delta_p u - |\nabla u|^q$$

- Modelos importantes de tipo "pilas de arena" (sandpiles) y una "fenomenología" matemática muy interesante y sorprendente motivan estos estudios.

# El flujo de la variación total

- El flujo de la variación total (total variation flow) se refiere a la ecuación 1-Laplaciana (es decir, la ecuación  $p$ -Laplaciana con  $p = 1$ ), que se escribe como

$$\partial_t u = \operatorname{div} \left( \frac{\nabla u}{|\nabla u|} \right)$$

- Vicent Caselles y (muchos) colaboradores han propuesto algoritmos de procesamiento de imágenes basados en las propiedades de esta ecuación, que se han vuelto muy usados en la práctica corriente (imágenes médicas, recuperación de imágenes borrosas o con defectos etc.), por ello es una ecuación que se ha vuelto muy famosa.
- El estudio matemático es muy complicado técnicamente debido a su singularidad: no regulariza "esquinas" (es decir, por ejemplo funciones características de una figura con esquinas, como una imagen) y sus soluciones son muy débiles.
- También en imágenes se usan modelos basados en ecuaciones de difusión no lineal todavía más singulares (Perona-Malik, PME con  $m < 0$  etc.).

# Los principales problemas que se estudian

# Existencia, unicidad y regularidad de soluciones

- Se trata de considerar el problema de Cauchy (para la PME, PLE u otra ecuación) con un dato inicial  $u_0$  con ciertas propiedades  $u_0 \in L^1(\mathbf{R}^N)$  es lo más habitual, también se suelen elegir datos iniciales más regulares (continuas, de clase  $C^1$ , solamente acotadas etc.) o el problema de Cauchy-Dirichlet en dominios acotados con una condición de frontera.
- Se quiere demostrar que dado un dato inicial, existe una solución de la ecuación  $u(x, t)$  (al menos para un corto intervalo de tiempo  $t \in (0, T)$ ) tal que  $u(x, 0) = u_0(x)$  para todo  $x$ .
- Habitualmente tenemos que definir el sentido de las soluciones (débiles, fuertes, clásicas, de viscosidad etc.). Eso es porque en general, las soluciones **no admiten** en todos los puntos **tantas derivadas** como para satisfacer la ecuación en sentido clásico.
- Ejemplo: las soluciones ZKB de la PME no son derivables hasta orden 2 en los puntos de interfase (frontera libre), por tanto en sentido clásico, en este punto **no se puede calcular**  $\Delta u^m$ . Por tanto, hay que introducir conceptos funcionales de soluciones (débiles, muy débiles, fuertes, de viscosidad, mild etc.) y tratar de probar la existencia y unicidad en estas clases de soluciones.

# Regularidad de soluciones

- El problema de la **regularidad de soluciones** significa establecer si, empezando con un dato inicial  $u_0$  en cierto espacio funcional, la función  $u(t)$  para todo tiempo  $t > 0$  mejora en cuanto a regularidad. Por ejemplo: si  $u_0 \in L^1(\mathbf{R}^N)$  pero  $u(t) \in C^2(\mathbf{R}^N)$ , la solución a tiempos posteriores ha mejorado mucho (desde una función solamente integrable hasta 2 veces derivable).
- Por ejemplo, en el caso de la PLE, se prueba que si  $u_0 \in L^1(\mathbf{R}^N)$ , entonces las soluciones a tiempo  $t > 0$  son de clase  $C^{1,\alpha}$ , es decir, derivables hasta orden 1 con derivadas parciales continuas en el sentido Holder.
- Los teoremas de regularidad suelen ser muy complicados técnicamente, pero muy útiles para poder trabajar con ciertas clases de soluciones.

## Comportamiento a largo plazo

- Es el problema probablemente más interesante y estudiado en relación a las ecuaciones de difusión no lineal. Necesitamos hacer "predicciones" sobre la forma y el recorrido que va a tener la difusión del gas, fluido etc. cuando dejamos pasar el tiempo. En matemáticas, se trata de estudiar el siguiente límite

$$\lim_{t \rightarrow \infty} t^\gamma \|u(t) - U(t)\|_\infty,$$

donde  $U(t)$  es un perfil (modelo) especial al que las soluciones se aproximan (y también, si es posible, encontrar de la forma más precisa posible la velocidad de convergencia).

- Por ejemplo, en el caso de la PME, las soluciones ZKB son los perfiles asintóticos de las soluciones

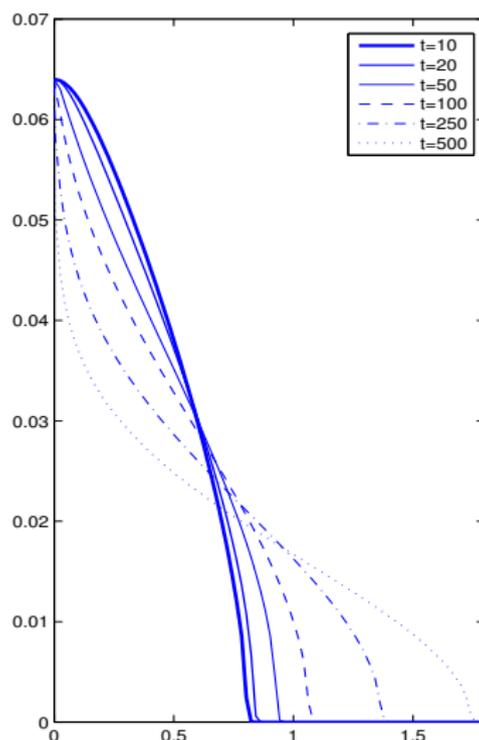
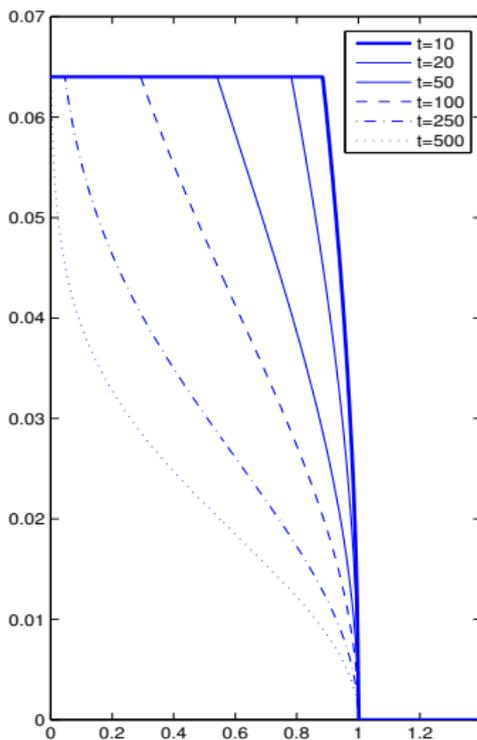
$$\lim_{t \rightarrow \infty} t^\alpha |u(x, t) - B(x, t)| = 0,$$

con convergencia uniforme, donde  $B(x, t)$  es la única solución ZKB cuya masa total (que se conserva) es igual a la masa del dato inicial  $u_0$ .

- Las siguientes figuras muestran la convergencia asintótica para la ecuación de los medios porosos no homogénea crítica

$$|x|^{-2} u_t = \Delta u^m, \quad m > 1$$

# Comportamiento a largo plazo



# Estudio de soluciones especiales

- En relación al comportamiento a largo plazo, se deben primero encontrar las soluciones "candidatas" a ser perfiles de comportamientos.
- Suelen tener ciertas simetrías especiales, y en las ecuaciones de difusión no lineal es habitual que tengan **forma auto-semejante**

$$U(x, t) = t^{-\alpha} f(|x|t^{-\beta}), \quad \text{forward self - similarity,}$$

o

$$U(x, t) = (T - t)^{\alpha} f(|x|(T - t)^{\beta}), \quad \text{backward self - similarity,}$$

y en algunos casos (bastante interesantes)

$$U(x, t) = e^{\alpha t} f(|x|e^{\beta t}), \quad \text{exponential self - similarity.}$$

- En muchos casos (sobre todo de ecuaciones "competitivas" o de difusión rápida) la clasificación de los perfiles (en forma auto-semejante) de comportamientos asintóticos es un problema muy complejo y serio, en cuyo estudio se utilizan **técnicas de sistemas dinámicos**, ya que muchos perfiles no son explícitos.

## Explosión o extinción en tiempo finito

- Finalmente, un problema muy importante en las ecuaciones "competitivas" de reacción-difusión o absorción-difusión es ver si pasan los siguientes fenómenos sobre las soluciones.
- **Explosión en tiempo finito:** partimos con  $u_0$  acotada, y existe  $T \in (0, \infty)$  tal que  $u(T)$  ya no está acotada en este tiempo: tiende a infinito en al menos un punto. Aparece con frecuencia en las ecuaciones de reacción-difusión: por ejemplo se sabe que **todas** las soluciones de

$$u_t = \Delta u^m + u^p, \quad m < p < m + \frac{2}{N}$$

explotan en tiempo finito.

- **Extinción en tiempo finito:** la propiedad de que una solución con dato inicial  $u_0$  positivo (al menos en cierto conjunto) se cancele en un tiempo finito  $T > 0$  (es decir  $u(T) \equiv 0$ ). Aparece con cierta frecuencia en el caso de ecuaciones de reacción-absorción, pero también, por ejemplo, se sabe que las soluciones de la PME con  $0 < m < m_c = (N - 2)/N$  tienen extinción, debido a la difusión muy rápida.
- En estos casos, se quiere estudiar cuándo surgen (explosión o extinción), en qué conjuntos, con qué tasa y con qué perfiles (suelen ser auto-semejantes de tipo "backward"). Son problemas bastante complicados pero muy interesantes.

FIN

Muchísimas gracias por  
vuestra asistencia.

# Los sonidos de $\pi$

Clara Simon de Blas  
Julio Flores Álvarez

# Los sonidos de $\pi$

¿Que pasa si asociamos números a las teclas de un piano?

David Macdonald ideo una manera de asociar números a las teclas de un piano para interpretar melodías con números.

# Los sonidos de $\pi$

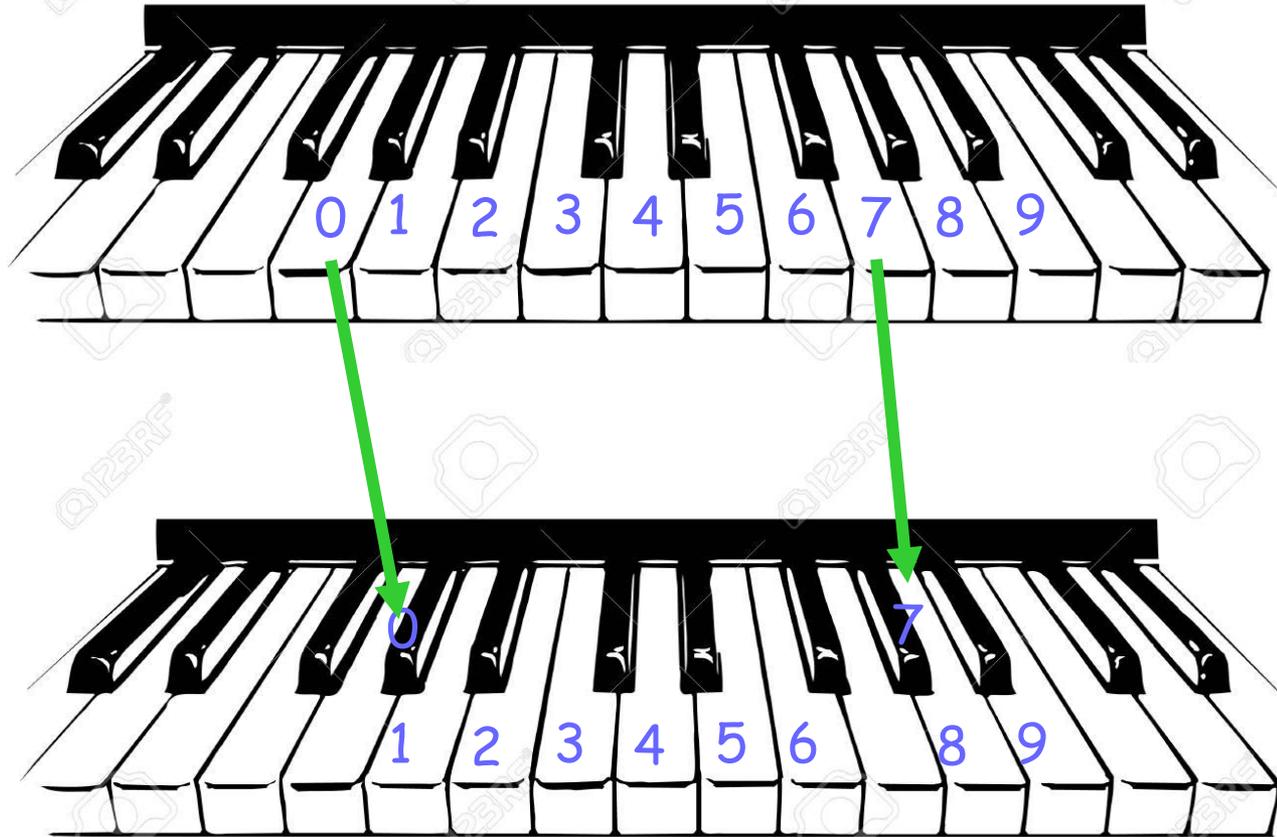


# Los sonidos de $\pi$

## Escala menor armónica en La

Es una escala menor a la cual le subimos el séptimo grado medio tono (en nuestro caso Sol) transformando la escala menor.

# Los sonidos de $\pi$



# Los sonidos de $\pi$

## Los armónicos de la mano izquierda

Nota Principal Nota Principal Nota Principal Nota Principal Nota Principal Nota Principal Nota Principal

Nota de Paso Nota de Paso Nota de Paso

Do Mayor Fa Mayor Sol Mayor Do Mayor

# Los sonidos de $\pi$

## Los armónicos de la mano izquierda

Do Mayor



Re Menor



Mi Menor



Fa Mayor



Sol Mayor



La Menor



Si Disminuido



# Los sonidos de $\pi$

## Los armónicos de la mano izquierda

Pno. *mf* 3. 1 4 1 5 9 2 6

The image shows a musical score for piano (Pno.) in treble clef. The score consists of two staves. The upper staff contains a sequence of chords, and the lower staff contains a sequence of eighth notes. The chords are indicated by the numbers 3, 1, 4, 1, 5, 9, 2, 6. The dynamic marking is *mf*. The score is labeled 'Pno.' on the left.

# Los sonidos de $\pi$

## Algunas curiosidades

- ❑ Su origen se remonta al año 2000 a.C
- ❑ Se emplea la letra griega pi (la decimosexta letra del alfabeto griego)
- ❑ Leonhard Euler, hacia 1734, fue el primero en saber su valor
- ❑ Arquímedes fue uno de los primeros en aproximarse al valor del número pi.
- ❑ William Shanks consiguió obtener 707 decimales del número pi tras un trabajo de investigación de casi 20 años.

# Los sonidos de $\pi$

## Algunas curiosidades

- ❑ Pegasus (1957) para calcular decimales de pi
- ❑ IBM 7090 (1961) logró llegar a 100.000 decimales de pi.
- ❑ 1999, Yasumasa Kanada y Daisuke Takahashi logran 206.158.430.000 decimales del número pi
- ❑ Se aplica a la fabricación de neumáticos, botellas, vasos o relojes.
- ❑ En astronomía, para calcular la cantidad de hidrógeno que se requiere en las misiones espaciales o para calcular las extensiones de territorio de los diferentes planetas.

# Los sonidos de $\pi$

Ya podemos interpretar  $\pi$

Nuestros artistas

Nicolás Flores

Julio Alberto Flores