

Congestión y penalización en transporte óptimo

Marcelo Gallardo

PUCP

10 de agosto de 2025

Introducción

El modelo

Propiedades del problema de optimización

Algoritmo

Estática comparativa

Aplicaciones

Extensiones

Referencias

Introducción

Introducción

- ▶ Matching [Gale and Shapley, 1962, Hylland and Zeckhauser, 1979, Kelso and Crawford, 1982, Roth, 1982, Roth and Sotomayor, 1990, Roth, 1982, Abdulkadiroğlu and Sönmez, 2003, Echenique and Yenmez, 2015, Echenique et al., 2023]
- ▶ Transporte óptimo [Villani, 2009, Galichon, 2016, Galichon, 2021, Dupuy and Galichon, 2022, Carlier et al., 2020].
- ▶ Gallardo, Loaiza & Chávez (2025). Comentarios de los profesores: Amilcar Velez, César Martinelli y Federico Echenique.

Matching

- ▶ Dos grupos: X e Y .
- ▶ Matching/asignación π .
- ▶ Caso discreto $X = \{x_1, \dots, x_N\}$, $Y = \{y_1, \dots, y_L\}$,
 $I = \{1, \dots, N\}$ y $J = \{1, \dots, L\}$ los índices.

$$\min_{\pi \in \Pi(\mu, \nu)} \sum_{i=1}^N \sum_{j=1}^L c_{ij} \pi_{ij}$$

con

$$\Pi(\mu, \nu) = \left\{ \pi_{ij} \geq 0 : \sum_{j=1}^L \pi_{ij} = \mu_i \quad \forall i \in I, \quad \sum_{i=1}^N \pi_{ij} = \nu_j \quad \forall j \in J \right\}. \quad (1)$$

Extensiones: regularización entrópica

$$\mathcal{P}_\varepsilon : \quad \min_{\pi \in \Pi(\mu, \nu)} \sum_{i=1}^N \sum_{j=1}^L c_{ij} \pi_{ij} - \varepsilon H(\pi), \quad (2)$$

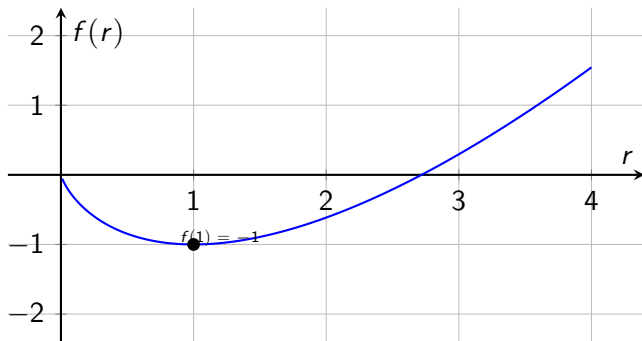
donde

$$H(\pi) = - \sum_{i,j} h(\pi_{i,j}) \quad (3)$$

$$h(r) = \begin{cases} r(\ln(r) - 1), & \text{if } r > 0 \\ 0, & \text{if } r = 0 \\ +\infty, & \text{if } r < 0. \end{cases}$$

π^* más densa.

Gráfica de $f(r) = r \ln r - r$



Modelo

Congestión y penalización

- Penalización: [Izmailov and Solodov, 2023].
- Congestión: MAE-CON 2.

$$\mathcal{P}_{CP} : \min_{\pi_{ij} \geq 0} \underbrace{\alpha \sum_{i=1}^N \sum_{j=1}^L \varphi(\pi_{ij}; \theta_{ij})}_{\text{Costo directo del matching}} + \underbrace{(1 - \alpha) \left[\sum_{i=1}^N \epsilon_i \left(\sum_{j=1}^L \pi_{ij} - \mu_i \right)^2 + \sum_{j=1}^L \delta_j \left(\sum_{i=1}^N \pi_{ij} - \nu_j \right)^2 \right]}_{\text{Costos de los objetivos sociales}}$$
$$=: F(\pi; \theta, \alpha, \epsilon, \delta, \mu, \nu) \quad (4)$$

donde $\epsilon_1, \dots, \epsilon_N$, $\delta_1, \dots, \delta_L$ y μ_1, \dots, μ_N , ν_1, \dots, ν_L son todos no negativos, y

$$\varphi(\pi_{ij}; \theta_{ij}) = d_{ij} + c_{ij}\pi_{ij} + a_{ij}\pi_{ij}^2. \quad (5)$$

Elementos del modelo

- ▶ $\sum_{\substack{1 \leq i \leq N \\ 1 \leq j \leq L}} c_{ij} \pi_{ij}$
- ▶ $\sum_{\substack{1 \leq i \leq N \\ 1 \leq j \leq L}} a_{ij} \pi_{ij}^2$
- ▶ μ_1, \dots, μ_N
- ▶ ν_1, \dots, ν_L
- ▶ $\epsilon_1, \dots, \epsilon_N$
- ▶ $\delta_1, \dots, \delta_L$
- ▶ $\alpha \in (0, 1), \alpha \rightarrow 1/2$

$$\sum_{j=1}^L \pi_{ij} = \mu_i \quad \Rightarrow \quad \epsilon_i \left[\sum_{j=1}^L \pi_{ij} - \mu_i \right]^2, \quad \epsilon_i > 0,$$

$$\sum_{i=1}^N \pi_{ij} = \nu_j \quad \Rightarrow \quad \delta_j \left[\sum_{i=1}^N \pi_{ij} - \nu_j \right]^2, \quad \delta_j > 0.$$

Interpretación

- ▶ Excesos de demanda u oferta.
- ▶ Mayor flexibilidad: α, ϵ, δ .
- ▶ Más sencillez en la resolución (CPO).

Nota: voy a tratar de manera indistinta $\mathcal{M}_{N \times L}$ y \mathbb{R}^{NL} usando el isomorfismo (vectorizando por filas no columnas).

Propiedades del problema

Propiedades

- ▶ Existencia y unicidad de la solución.
- ▶ Solución en (F coerciva):

$$\Omega = [0, R]^{NL}, \text{ with } R = N \max_{1 \leq i \leq N} \{\mu_i\} + L \max_{1 \leq j \leq L} \{\nu_j\}.$$

- ▶ Optimización sobre \mathbb{R}_+^{NL} :
 - ▶ Existen cotas de proximidad entre soluciones enteras y continuas para programas cuadráticos convexos separables con restricciones lineales [Hochbaum and Shanthikumar, 1990].
 - ▶ [Granot and Skorin-Kapov, 1990]: para una solución entera óptima $\pi_{\mathbb{Z}}^*$, existe $\pi_{\mathbb{R}}^*$ tal que:

$$\|\pi_{\mathbb{Z}}^* - \pi_{\mathbb{R}}^*\|_{\infty} \leq n \cdot \Delta(\Lambda), \text{ con (en nuestro caso) } \underbrace{\Delta(\Lambda)}_{\det(\text{Restricciones})} = 1.$$

- ▶ Ej.: [Park and Boyd, 2017] muestran que redondear $\pi_{\mathbb{R}}^*$ genera un error acotado: $|F(\pi_{\mathbb{Z}}^*) - F(\pi_{\mathbb{R}}^*)| \leq \frac{NL}{4} \cdot \lambda_{\max}$, donde λ_{\max} es el mayor autovalor del Hessiano de F .

Condiciones de primer orden

$$\begin{aligned} \frac{\partial F}{\partial \pi_{ij}} = \frac{1}{2} & \left(\varphi'(\pi_{ij}^*; \theta_{ij}) + 2\epsilon_i \left(\sum_{\ell=1}^L \pi_{i\ell}^* - \mu_i \right) \right. \\ & \left. + 2\delta_j \left(\sum_{k=1}^N \pi_{kj}^* - \nu_j \right) - \gamma_{ij}^* \right) = 0, \quad \forall (i, j) \in I \times J. \end{aligned}$$

- ▶ γ_{ij} : multiplicador asociado a la condición de no negatividad.
- ▶ En el caso $\gamma_{ij}^* = 0$, se cumple que $\nabla F(\pi^*) = 0$.
- ▶ Esto implica $A\pi^* = b$, donde:

$$\begin{aligned} A = & \underbrace{\text{Diag}(a_{11}, a_{12}, \dots, a_{NL})}_D + \underbrace{\text{Diag}(\epsilon_1, \dots, \epsilon_N) \otimes \mathbf{1}_{L \times L}}_E \\ & + \underbrace{\mathbf{1}_{N \times N} \otimes \text{Diag}(\delta_1, \dots, \delta_L)}_F, \end{aligned}$$

$$b = \left[\epsilon_1 \mu_1 + \delta_1 \nu_1 - \frac{c_{11}}{2}, \dots, \epsilon_N \mu_N + \delta_L \nu_L - \frac{c_{NL}}{2} \right]^T.$$

$$D = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{12} & \cdots & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & a_{NL} \end{bmatrix}$$

$$E = \begin{bmatrix} \epsilon_1 & 0 & \cdots & 0 \\ 0 & \epsilon_2 & \cdots & \vdots \\ \vdots & & \ddots & \\ 0 & & \cdots & \epsilon_N \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & \vdots \\ \vdots & & \ddots & \\ 1 & \cdots & & 1 \end{bmatrix}$$

$$F = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & \vdots \\ \vdots & & \ddots & \\ 1 & \cdots & & 1 \end{bmatrix} \otimes \begin{bmatrix} \delta_1 & 0 & \cdots & 0 \\ 0 & \delta_2 & \cdots & \vdots \\ \vdots & & \ddots & \\ 0 & & \cdots & \delta_L \end{bmatrix}.$$

Ejemplo $N = L = 2$

$$A = \begin{bmatrix} a_{11} + \epsilon_1 + \delta_1 & \epsilon_1 & \delta_1 & 0 \\ \epsilon_1 & a_{12} + \epsilon_1 + \delta_2 & 0 & \delta_2 \\ \delta_1 & 0 & a_{21} + \epsilon_2 + \delta_1 & \epsilon_2 \\ 0 & \delta_2 & \epsilon_2 & a_{22} + \epsilon_2 + \delta_2 \end{bmatrix}$$

Producto Kronecker

Dadas $A \in \mathcal{M}_{m \times n}$ y $B \in \mathcal{M}_{p \times q}$

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \in \mathcal{M}_{mp \times nq}$$

Si A y B son matrices cuadradas de tamaños respectivos n y q .
Sean $\lambda_1, \dots, \lambda_n$ los autovalores de A y μ_1, \dots, μ_q los de B .
Entonces, los autovalores de $A \otimes B$ son:

$$\lambda_i \mu_j, \quad i = 1, \dots, n, \quad j = 1, \dots, q.$$

Determinante de A

Lema

El determinante de A es estrictamente positivo si todos los parámetros son estrictamente positivos.

- ▶ $\det(D) = \prod_{(i,j)} a_{ij} > 0$, ya que D es diagonal positiva.
- ▶ $\det(E) = \det(F) = 0$, pero ambas matrices son simétricas y semidefinidas positivas.
- ▶ Entonces, $A = D + E + F$ es suma de una matriz diagonal positiva y dos matrices simétricas semidefinidas positivas.
- ▶ Por la desigualdad de Minkowski para determinantes, [Artstein-Avidan et al., 2015], [Zhan, 2005]:

$$\det(A) \geq \det(D + E) + \det(F) \geq \det(D) > 0.$$

Resultados

Invertir A

- ▶ Se puede.
- ▶ Fórmula cerrada: complicado (Wolfram 3×3).
- ▶ Aproximación de Neumann.
- ▶ Sherman-Morrison: Algoritmo para invertir.

Convergencia y desarrollo en serie de A^{-1}

Supuesto

Supóngase que $a_{ij} > 0$ y que:

$$\max_i \{\epsilon_i\} \cdot L + \max_j \{\delta_j\} \cdot N < \min_{(i,j)} \{a_{ij}\}.$$

- ▶ Esta condición garantiza que $\|D^{-1}X\|_\sigma < 1$ (norma espectral).
- ▶ Bajo esta condición:

$$A^{-1} = \left(\sum_{k=0}^{\infty} (-1)^k (D^{-1}X)^k \right) D^{-1}.$$

- ▶ Se obtiene convergencia del método iterativo:

$$\pi_n \rightarrow \pi^* = A^{-1}b.$$

Prueba de la expansión

- ▶ Análisis en EVN dimensión finita.
- ▶ $A = D + E + F = D + X$.
- ▶ $A^{-1} = (D + X)^{-1} = (I - (-1)D^{-1}X)^{-1}D^{-1}$.
- ▶ $\forall \lambda \in \sigma(D^{-1}X)$,

$$\lambda \leq \max_{i,j} \{1/a_{ij}\} \cdot (\lambda_{\max}^E + \lambda_{\max}^F)$$

donde $\lambda_{\max}^E = \max_i \{\epsilon_i\} \cdot L$ y $\lambda_{\max}^F = \max_j \{\delta_j\} \cdot N$.

- ▶ Por ende, $\|D^{-1}X\|_{\sigma} < 1$.
- ▶ Podemos escribir

$$(I - (-1)D^{-1}X)^{-1} = \sum_{k=0}^{\infty} (-1)^k (D^{-1}X)^k.$$

Prueba de la convergencia

Claramente $\|\pi_n - \pi^*\|_\infty = \|\mathcal{E}_n b\|_\infty \leq \|\mathcal{E}_n b\|_2$

$$\begin{aligned}\|\mathcal{E}_n b\|_2 &\leq \sqrt{NL} \left\| \sum_{k=n+1}^{\infty} (-1)^k (D^{-1}X)^k \right\|_\sigma \|D^{-1}b\|_\infty \\ &\leq \frac{\sqrt{NL} \|D^{-1}X\|_\sigma^{n+1} \|D^{-1}b\|_\infty}{1 - \|D^{-1}X\|_\sigma}.\end{aligned}$$

Entonces, dado $\varepsilon > 0$,

$$N_\varepsilon = \max \left\{ 1, \left\lceil \log_{\|D^{-1}X\|_\sigma} \left(\frac{\varepsilon (1 - \|D^{-1}X\|_\sigma)}{\sqrt{NL} \|D^{-1}b\|_\infty} \right) \right\rceil + 1 \right\},$$

para $n \geq N_\varepsilon$, se sigue que $\|\pi_n - \pi^*\|_\infty < \varepsilon$.

Se tiene que

$$\frac{\sqrt{NL} \|D^{-1}X\|_{\sigma}^{n+1} \|D^{-1}b\|_{\infty}}{1 - \|D^{-1}X\|_{\sigma}} \leq \varepsilon$$

$$\sqrt{NL} \|D^{-1}X\|_{\sigma}^{n+1} \|D^{-1}b\|_{\infty} < \varepsilon(1 - \|D^{-1}X\|_{\sigma})$$

$$\|D^{-1}X\|_{\sigma}^{n+1} < \frac{\varepsilon(1 - \|D^{-1}X\|_{\sigma})}{\sqrt{NL} \|D^{-1}b\|_{\infty}}$$

$$n + 1 \geq \log_{\|D^{-1}X\|_{\sigma}} \left(\frac{\varepsilon(1 - \|D^{-1}X\|_{\sigma})}{\sqrt{NL} \|D^{-1}b\|_{\infty}} \right).$$

Caso especial: congestión homogénea y sin cupos

Supuesto

Supóngase que $\delta_j = 0$ para todo j , y que $D = \beta I$ con $\beta > 0$.

- Se tiene una fórmula explícita para A^{-1} :

$$A^{-1} = \frac{I}{\beta} + \frac{1}{\beta} \text{Diag} \left(-\frac{\epsilon_1}{\beta + L\epsilon_1}, \dots, -\frac{\epsilon_N}{\beta + L\epsilon_N} \right) \otimes \mathbf{1}_{L \times L}.$$

- Bajo esta expresión, se deduce una fórmula explícita para π_{ij}^* :

$$\pi_{ij}^* = \frac{b_{ij}}{\beta} - \sum_{\ell=1}^L \frac{b_{i\ell}\epsilon_i}{\beta^2 + L\epsilon_i\beta}.$$

Prueba

$$\begin{aligned} A^{-1} &= \frac{I}{\beta} + \left[\sum_{k=1}^{\infty} (-1)^k \left(\frac{1}{\beta} \right)^k (\text{Diag}(\epsilon_1, \dots, \epsilon_N) \otimes \mathbf{1}_{L \times L})^k \right] D^{-1} \\ &= \frac{I}{\beta} + \frac{1}{\beta L} \sum_{k=1}^{\infty} (-1)^k \left(\frac{L}{\beta} \right)^k (\text{Diag}(\epsilon_1^k, \dots, \epsilon_N^k) \otimes \mathbf{1}_{L \times L}) \\ &= \frac{I}{\beta} + \frac{1}{\beta L} \text{Diag} \left(\sum_{k=1}^{\infty} (-1)^k \left(\frac{L\epsilon_1}{\beta} \right)^k, \dots, \sum_{k=1}^{\infty} (-1)^k \left(\frac{L\epsilon_N}{\beta} \right)^k \right) \otimes \mathbf{1}_{L \times L} \\ &= \frac{I}{\beta} + \frac{1}{\beta} \text{Diag} \left(-\frac{\epsilon_1}{\beta + L\epsilon_1}, \dots, -\frac{\epsilon_N}{\beta + L\epsilon_N} \right) \otimes \mathbf{1}_{L \times L}. \end{aligned}$$

Usamos la propiedades del producto Kronecker:

- ▶ $(A \otimes B)^k = A^k \otimes B^k$.
- ▶ $\mathbf{1}_{L \times L}^k = L^{k-1} \mathbf{1}_{L \times L}$.

Caso especial: pesos homogéneos

Supuesto

Sean $a_{ij} = \rho$ y $\epsilon_i = \delta_j = \zeta$ para todo (i, j) , con $\rho > 2NL\zeta > 0$.

- ▶ En este caso, $A = \rho I + \zeta Y$ donde Y es una matriz estructurada.

- ▶ Entonces:

$$A^{-1} = \frac{1}{\rho} \sum_{k=0}^{\infty} \left(-\frac{\zeta}{\rho} \right)^k Y^k.$$

- ▶ Se pueden acotar los términos $(A^{-1})_{ij}$ por constantes explícitas C_1, C_2 .
- ▶ Concretamente:

$$\pi_{ij}^* \leq NL \cdot \max\{|C_1|, C_2\} \cdot \max_{i,j} \left| \epsilon_i \mu_i + \delta_j \nu_j - \frac{c_{ij}}{2} \right|.$$

Donde:

$$C_1 = \frac{\zeta \left(4\zeta N^3 L^3 (2\zeta^3 - 2\zeta \rho^2 - \rho^3) + 8N^2 L^2 \rho^2 (\rho^2 - \zeta^2) + \zeta NL \rho^2 (2\zeta + \rho) - 2\rho^4 \right)}{\rho^4 (\zeta^2 NL - \rho^2) (2NL - 1) (2NL + 1)}$$

$$C_2 = \frac{\zeta^2 NL \rho (4NL - 1)}{(\rho^2 - \zeta^2 NL) (\rho - 2NL\zeta) (\rho + 2NL\zeta)}.$$

Algoritmo

Algunos preliminares

Primero, escribamos

$$\begin{aligned} A = \text{Diag}(a_{11}, \dots, a_{NL}) &+ \sum_{i=1}^N \left(\epsilon_i^{1/2} \mathbf{e}_i \otimes \mathbf{1}_{L \times 1} \right) \left(\epsilon_i^{1/2} \mathbf{e}_i^T \otimes \mathbf{1}_{1 \times L} \right) \\ &+ \sum_{j=1}^L \left(\delta_j^{1/2} \mathbf{1}_{N \times 1} \otimes \mathbf{e}_j \right) \left(\delta_j^{1/2} \mathbf{1}_{1 \times N} \otimes \mathbf{e}_j^T \right). \end{aligned} \quad (6)$$

Usamos la propiedad del producto de Kronecker:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD),$$

siempre que las dimensiones sean compatibles. Entonces, para cada i se tiene:

$$\begin{aligned} \left(\epsilon_i^{1/2} \mathbf{e}_i \otimes \mathbf{1}_{L \times 1} \right) \left(\epsilon_i^{1/2} \mathbf{e}_i^\top \otimes \mathbf{1}_{1 \times L} \right) &= \epsilon_i \left(\mathbf{e}_i \mathbf{e}_i^\top \otimes \mathbf{1}_{L \times 1} \cdot \mathbf{1}_{1 \times L} \right) \\ &= \epsilon_i \left(\mathbf{e}_i \mathbf{e}_i^\top \otimes \mathbf{1}_{L \times L} \right). \end{aligned}$$

Sumando sobre $i = 1, \dots, N$:

$$\begin{aligned} \sum_{i=1}^N \epsilon_i \left(\mathbf{e}_i \mathbf{e}_i^\top \otimes \mathbf{1}_{L \times L} \right) &= \left(\sum_{i=1}^N \epsilon_i \mathbf{e}_i \mathbf{e}_i^\top \right) \otimes \mathbf{1}_{L \times L} \\ &= \text{Diag}(\epsilon_1, \dots, \epsilon_N) \otimes \mathbf{1}_{L \times L}. \end{aligned}$$

Motivación

- ▶ Supongamos que ya conocemos la inversa de una matriz A^{-1} .
- ▶ Recibe un cambio pequeño: una actualización de rango 1 de la forma $A + uv^T$.
- ▶ Notar que

$$uv^T = \begin{bmatrix} v_1 u & \cdots & v_n u \end{bmatrix}.$$

- ▶ ¿Debemos volver a invertir toda la matriz?
- ▶ ¡No! Usemos Sherman–Morrison.

Sherman–Morrison

Sean:

- ▶ $A \in \mathbb{R}^{n \times n}$: matriz invertible,
- ▶ $u, v \in \mathbb{R}^n$: vectores columna,
- ▶ uv^\top : producto exterior (rango 1).

Si $1 + v^\top A^{-1}u \neq 0$, entonces:

$$(A + uv^\top)^{-1} = A^{-1} - \frac{A^{-1}uv^\top A^{-1}}{1 + v^\top A^{-1}u}$$

Actualizaciones de rango 1

Los siguientes términos son **actualizaciones de rango 1**:

$$\epsilon_i^{1/2} \mathbf{e}_i \otimes \mathbf{1}_{L \times 1} \cdot \epsilon_i^{1/2} \mathbf{e}_i^\top \otimes \mathbf{1}_{1 \times L} \quad \text{y} \quad \delta_j^{1/2} \mathbf{1}_{N \times 1} \otimes \mathbf{e}_j \cdot \delta_j^{1/2} \mathbf{1}_{1 \times N} \otimes \mathbf{e}_j^\top.$$

Usando las propiedades del producto Kronecker
 $((A \otimes B)^\top = A^\top \otimes B^\top)$:

$$\epsilon_i \left(\mathbf{e}_i \mathbf{e}_i^\top \otimes \mathbf{1}_{L \times L} \right) \quad \text{y} \quad \delta_j \left(\mathbf{1}_{N \times N} \otimes \mathbf{e}_j \mathbf{e}_j^\top \right).$$

$$N = 3 \text{ y } L = 2$$

Primero:

$$e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n \Rightarrow e_2 e_2^\top = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

Luego:

$$\begin{aligned} \epsilon_2 \left(e_2 e_2^\top \otimes \mathbf{1}_{2 \times 2} \right) &= \epsilon_2 \cdot \begin{bmatrix} \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \mathbf{1}_{2 \times 2} & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 2} \end{bmatrix} \\ &= \epsilon_2 \cdot \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

¿Quién es u en Sherman–Morrison?

En cada paso, aplicamos Sherman–Morrison para una matriz de la forma $A + uu^\top$.

- ▶ En el bucle sobre $i = 1, \dots, N$:

$$u^{(i)} := \epsilon_i^{1/2} e_i \otimes \mathbf{1}_{L \times 1}$$

- ▶ En el bucle sobre $j = 1, \dots, L$:

$$v^{(j)} := \delta_j^{1/2} \mathbf{1}_{N \times 1} \otimes e_j$$

Ambos vectores definen productos exteriores del tipo uv^\top , que son actualizaciones de rango 1.

Aplicamos:

$$A^{-1} \leftarrow A^{-1} - \frac{A^{-1} u u^\top A^{-1}}{1 + u^\top A^{-1} u}$$

Condición clave: $1 + u^\top A^{-1} u \neq 0$

Para que la fórmula de Sherman–Morrison sea válida, es necesario que:

$$1 + u^\top A^{-1} u \neq 0$$

- ▶ Esto se cumple automáticamente si:
 - ▶ $A^{-1} \succ 0$ (matriz definida positiva),
 - ▶ $u \neq 0$,
- ▶ En tu caso:
 - ▶ Se inicializa $A^{-1} = \text{Diag}(1/a_{ii})$, con $a_{ii} > 0$,
 - ▶ Los vectores $u^{(i)}$ y $v^{(j)}$ son no nulos porque $\epsilon_i, \delta_j > 0$.

Entonces: todas las actualizaciones cumplen la condición.

Algorithm: OPTIMIZE($a, b, \epsilon_1, \dots, \epsilon_N, \delta_1, \dots, \delta_L$)

- 1: **Input:** Vector $a \in \mathbb{R}_{++}^{NL}$, vector $b \in \mathbb{R}^{NL}$ and parameters $\epsilon_1, \dots, \epsilon_N, \delta_1, \dots, \delta_L$ in \mathbb{R}_{++}
- 2: **Output:** $\pi^* \in \mathbb{R}^{NL}$
- 3: Initialize $A^{-1} \leftarrow \text{Diag}(1/a_{11}, \dots, 1/a_{NL})$
- 4: **for** $i = 1$ to N **do**
- 5: Define $u^{(i)} \leftarrow \epsilon_i^{1/2} e_i \otimes 1_{L \times 1}$
- 6: $A^{-1} \leftarrow A^{-1} - \frac{A^{-1} u^{(i)} u^{(i)T} A^{-1}}{1 + u^{(i)T} A^{-1} u^{(i)}}$
- 7: **end for**
- 8: **for** $j = 1$ to L **do**
- 9: Define $v^{(j)} \leftarrow \delta_j^{1/2} 1_{N \times 1} \otimes e_j$
- 10: $A^{-1} \leftarrow A^{-1} - \frac{A^{-1} v^{(j)} v^{(j)T} A^{-1}}{1 + v^{(j)T} A^{-1} v^{(j)}}$
- 11: **end for**
- 12: **return** $A^{-1} b$

Teorema: Complejidad del Algoritmo OPTIMIZE

Theorem

Sea π^ una solución interior del sistema $A\pi = b$. Entonces, el algoritmo OPTIMIZE computa π^* en tiempo*

$$O((N + L)(NL)^2).$$

Idea central de la demostración

- ▶ La matriz A tiene la forma:

$$A = \text{Diag}(a) + \sum_{i=1}^N u^{(i)} u^{(i)\top} + \sum_{j=1}^L v^{(j)} v^{(j)\top}$$

donde $u^{(i)}$ y $v^{(j)}$ son vectores definidos por productos de Kronecker.

- ▶ Se inicializa $A^{-1} = \text{Diag}(1/a)$, y se aplican sucesivas actualizaciones de rango 1.
- ▶ Cada actualización usa la fórmula de Sherman–Morrison:

$$A^{-1} \leftarrow A^{-1} - \frac{A^{-1} w w^{\top} A^{-1}}{1 + w^{\top} A^{-1} w},$$

donde $w = u^{(i)}$ o $w = v^{(j)}$.

Cálculo de la complejidad

- ▶ Hay $N + L$ actualizaciones.
- ▶ Cada actualización requiere computar:
 - ▶ $w^\top A^{-1}w$: $O((NL)^2)$,
 - ▶ $A^{-1}w$: $O((NL)^2)$, $NL \times NL$ (se recorre toda la matriz)
 - ▶ $A^{-1}ww^\top A^{-1}$: $O((NL)^2)$ con una buena parentización (agrupar ww^\top perjudica).
- ▶ Así, en total: $O((N + L)(NL)^2)$.
- ▶ Manuel Loaiza Vasquez (Autodesk, Software Engineer).

Comparación de algoritmos para resolver $A\pi = b$

Tiempo	A dispersa	Galáctico	Referencia
$O((NL)^3)$	No	No	Folklore
$O((NL)^{2,81})$	No	No	Strassen (1969)
$O((N + L)(NL)^2)$	No	No	Este trabajo
$O((NL)^{2,371339})$	No	Sí	Alman (2025)
$O((NL)^{2,331645})$	Sí	Sí	PV (2024)

Cuadro: Comparación de algoritmos para resolver sistemas lineales.

Progreso reciente en resolución de sistemas

Resolver el sistema lineal $A\pi = b$ de forma eficiente ha sido objeto de avances importantes la última década para clases especiales de matrices:

- ▶ Sistemas simétricos diagonalmente dominantes $a_{ii} > \sum_{j \neq i} |a_{ij}|$ [Koutis et al., 2012]
- ▶ Casos donde A es dispersa [Peng and Vempala, 2024]: mayoría de entradas cero.

Muchos trabajos recientes se enfocan en algoritmos aproximados para computar A^{-1} eficientemente, o directamente resolver $Ax = b$. Sin embargo, estos enfoques:

- ▶ No aplican al caso general,
- ▶ Y algunos no son algoritmos exactos.

Inversión vs. multiplicación

- ▶ La inversión de matrices puede reducirse a multiplicación matricial en muchos algoritmos con tiempos asintóticamente equivalentes.
- ▶ Esta reducción se discute en los apuntes de clase de Vassilevska Williams: *CS367 Algebraic Graph Algorithms, Lectures 1 and 2*, Stanford University, 2015. Scribed by Jessica Su.
Disponible en
<https://theory.stanford.edu/~virgi/cs367/lecture1.pdf>

Métodos galácticos de multiplicación

- ▶ Alman et al. [Alman et al., 2025] presentan los mejores tiempos teóricos conocidos para multiplicación de matrices, basados en el *laser method* [Strassen, 1986], [Coppersmith and Winograd, 1990], [Davie and Stothers, 2013], [Williams, 2012].
- ▶ Pero estos algoritmos son actualmente imprácticos para su uso real.
- ▶ *Nota: estos son ejemplos de algoritmos galácticos*¹.

¹Término propuesto por R. Lipton para algoritmos asintóticamente superiores, pero inservibles en la práctica.

Implementación real: métodos Strassen-like

- ▶ Algoritmos como Strassen [Strassen, 1969], Winograd [Winograd, 1971], Pan [Pan, 1982], Karstadt [Karstadt and Schwartz, 2017] sí se han implementado, evaluado y usado en la práctica.
- ▶ Dependiendo del tamaño de matriz y constantes ocultas, son eficientes.
- ▶ Benchmarks reales muestran sus ventajas cuando el tamaño lo justifica [Huang, 2018].

Resultado

- ▶ Cuando $L \in \Theta(N)$, nuestro algoritmo alcanza la mejor cota superior conocida frente a cualquier método exacto derivado de multiplicación rápida.
- ▶ Aprovecha la estructura (diagonal + rango 1), no requiere aproximaciones.
- ▶ Computacionalmente estable y práctico.

Estática comparativa

Estática comparativa: aproximación con D^{-1}

- ▶ Aunque podemos computar π^* con la serie de Neumann o el Algoritmo 1, no es fácil obtener una expresión cerrada.
- ▶ Como primera aproximación, se propone usar $A^{-1} \approx D^{-1}$:

$$\pi_{ij}^* \simeq \frac{2(\epsilon_i \mu_i + \delta_j \nu_j) - c_{ij}}{2a_{ij}}.$$

- ▶ Derivadas parciales: $\partial \pi_{ij}^* / \partial a_{ij}, \partial \pi_{ij}^* / \partial c_{ij} < 0$;
 $\partial \pi_{ij}^* / \partial \epsilon_i, \partial \pi_{ij}^* / \partial \delta_j, \partial \pi_{ij}^* / \partial \mu_i, \partial \pi_{ij}^* / \partial \nu_j > 0$.
- ▶ Pero: $\partial \pi_{ij}^* / \partial \theta_{k\ell} = 0$ si $(k, \ell) \neq (i, j)$: no hay efecto sustitución.

Estática comparativa: aproximación refinada

- ▶ Usamos ahora $A^{-1} \approx D^{-1} - (D^{-1})^2 X$.
- ▶ Resultado clave:

$$\left[\frac{\partial \pi_{ij}^*}{\partial \theta_{k\ell}} \right] \simeq - \left[D^{-1} - (D^{-1})^2 X \mid A_{\Pi,2}^{-1} \right],$$

donde $A_{\Pi,2}^{-1}$ multiplica cada columna por π_{ij}^* .

- ▶ Si $\epsilon_i + \delta_j < 1$:
 - ▶ $\partial \pi_{ij}^* / \partial \theta_{ij} < 0$
 - ▶ $\partial \pi_{ij}^* / \partial \theta_{i\ell}, \partial \pi_{ij}^* / \partial \theta_{kj} > 0$
 - ▶ $\partial \pi_{ij}^* / \partial \theta_{k\ell} = 0$ si $i \neq k, j \neq \ell$

Estática comparativa: derivadas explícitas

$$\partial \pi_{ij}^* / \partial c_{ij} = - \frac{1 - (\epsilon_i + \delta_j)}{a_{ij}^2} < 0,$$

$$\partial \pi_{ij}^* / \partial c_{i\ell} = \frac{\epsilon_i}{a_{ij}^2} > 0, \quad \partial \pi_{ij}^* / \partial c_{kj} = \frac{\delta_j}{a_{ij}^2} > 0,$$

$$\partial \pi_{ij}^* / \partial c_{k\ell} = 0 \text{ si } i \neq k, j \neq \ell.$$

$$\partial \pi_{ij}^* / \partial a_{ij} = - \frac{2\pi_{ij}^*(1 - (\epsilon_i + \delta_j))}{a_{ij}^2} < 0,$$

$$\partial \pi_{ij}^* / \partial a_{i\ell} = \frac{2\pi_{i\ell}^* \epsilon_i}{a_{ij}^2} > 0, \quad \partial \pi_{ij}^* / \partial a_{kj} = \frac{2\pi_{kj}^* \delta_j}{a_{ij}^2} > 0,$$

$$\partial \pi_{ij}^* / \partial a_{k\ell} = 0 \text{ si } i \neq k, j \neq \ell.$$

Estática comparativa

- ▶ El efecto sustitución está ahora presente: si aumenta el costo de unir a i con j , π_{ij}^* baja y $\pi_{i\ell}^*$, π_{kj}^* suben.
- ▶ Aunque es una aproximación (Neumann truncada), capta la intuición económica.

Aplicaciones del modelo

Motivación: Aplicamos el modelo en sectores críticos del Perú donde existen problemas estructurales y costos de acceso significativos.

- ▶ El modelo es **normativo** (prescriptivo), no descriptivo.
- ▶ Permite capturar congestión, ineficiencias y demandas no satisfechas.
- ▶ Da flexibilidad al SP.

Salud

- ▶ Fragmentado en SIS (Seguro Integral de Salud), EsSalud, y EPS (Entidades Prestadoras de Salud) [Anaya-Montes and Gravelle, 2024].
 1. EPS: privado.
 2. EsSalud: trabajadores formales, tanto privado como público.
 3. SIS: universal, aplica a informales.
- ▶ SIS y EsSalud: $\simeq 80\%$. $\simeq 10\%$ EPS.
- ▶ Cobertura: sin intersecciones, salvo EPS y EsSalud [Velasquez, 2020].

Problema identificado	Indicador cuantificable
Escasez de personal médico en atención primaria	12 médicos por cada 10,000 habitantes, muy por debajo del estándar de 43 recomendado por la OMS [Bendezu-Quispe et al., 2020].
Falta de camas hospitalarias	1.6 camas por cada 1,000 habitantes, debajo del promedio regional [World Bank, 2020].
Congestión en UCI neonatales	50 % de las unidades presentan ineficiencia por sobrecarga de pacientes [Arrieta and Guillén, 2017].
Ineficiencias en el sistema de referencias	Alta proporción de pacientes tratados en centros no equipados [Soto, 2019].
Baja cobertura y largas esperas	Cobertura de hasta 86 % para ciertos tratamientos complejos [EsSalud, 2025a].
Altas tasas de postergación en ciudades clave	Más del 23 % de citas postergadas (enero–marzo 2025) [EsSalud, 2025b].

Nota: En 2016, el MINSA reportó un déficit de más de 47,000 profesionales de salud. El 36 % de los centros de mediana y alta complejidad carecía de personal suficiente, el 44 % de equipamiento y el 25 % de infraestructura adecuada.

Ejemplo

- ▶ 3 grupos
- ▶ Fricciones en los grupos: sectores económicos, geografía.
- ▶ Tabla (slide anterior): congestión, excesos de demanda.

Ejemplo (continúa)


- ▶ Resolvemos el problema \mathcal{P}_{CP} con $\mathbf{d} = 5\mathbf{I}_{3 \times 3}$ y $\alpha = 0,5$.
- ▶ Parámetros utilizados:

$$\mathbf{c} = \begin{bmatrix} 1 & 50 & 20 \\ 50 & 1 & 20 \\ 20 & 10 & 1 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 1 & 5 & 10 \\ 5 & 1 & 2 \\ 10 & 5 & 1 \end{bmatrix}, \quad \epsilon = \delta = \begin{bmatrix} 0,3 \\ 0,3 \\ 0,3 \end{bmatrix}$$

$$\mu = \begin{bmatrix} 100 \\ 50 \\ 20 \end{bmatrix}, \quad \nu = \begin{bmatrix} 90 \\ 40 \\ 40 \end{bmatrix}.$$

- ▶ Solución óptima π^* obtenida en Mathematica 14.1 usando el Algoritmo 1²:

$$\pi^* = \begin{bmatrix} 34,7802 & 0,19412 & 1,65935 \\ 0,10148 & 15,6978 & 3,41038 \\ 0,883807 & 0,905689 & 9,65139 \end{bmatrix}.$$

²También verificamos el resultado con `QuadraticOptimization`. 

Comparación

- ▶ Se resuelve \mathcal{P}_Q y \mathcal{P}_O con las mismas entradas que en \mathcal{P}_{CP} , pero imponiendo $\Pi(\boldsymbol{\mu}, \boldsymbol{\nu})$ y eliminando penalizaciones.
- ▶ Soluciones óptimas:

$$\boldsymbol{\pi}_{\mathcal{P}_Q}^* = \begin{bmatrix} 84,275 & 8,84062 & 6,88442 \\ 4,2985 & 30,4206 & 15,2809 \\ 1,42655 & 0,73873 & 17,8347 \end{bmatrix}, \quad \boldsymbol{\pi}_{\mathcal{P}_O}^* = \begin{bmatrix} 90 & 0 & 10 \\ 0 & 40 & 10 \\ 0 & 0 & 20 \end{bmatrix}.$$

Educación

- ▶ Tráfico vehicular.
- ▶ Preferencias: Afinidad académica-cultural, entorno social.
- ▶ Capacidad (instalación).
- ▶ Álvaro Mauricio Morán (usando el API de Google Maps).



Figura: Lunes 7pm

Científica del sur



Figura: Lunes 7pm



Figura: Lunes 7pm



Figura: Lunes 7pm



Figura: Lunes 7pm



Figura: Lunes 7pm

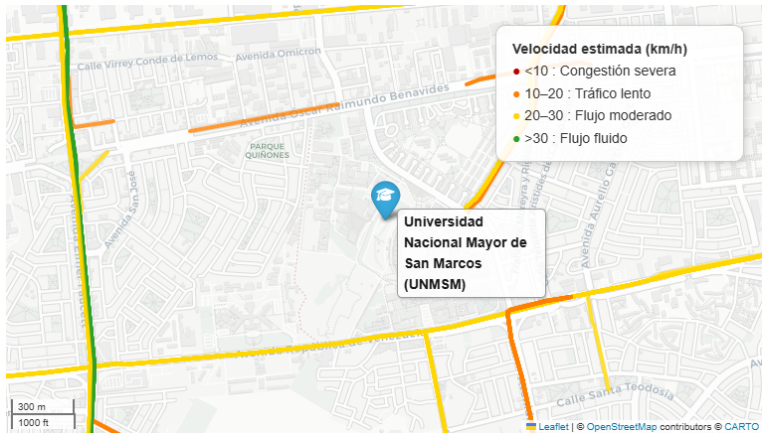


Figura: Lunes 7pm

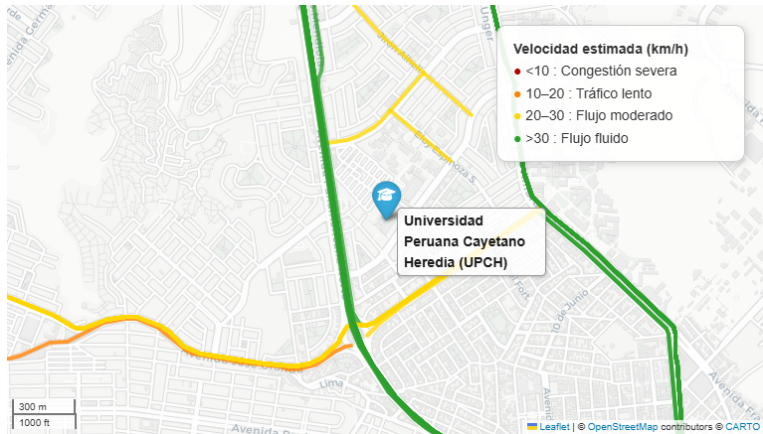


Figura: Lunes 7pm

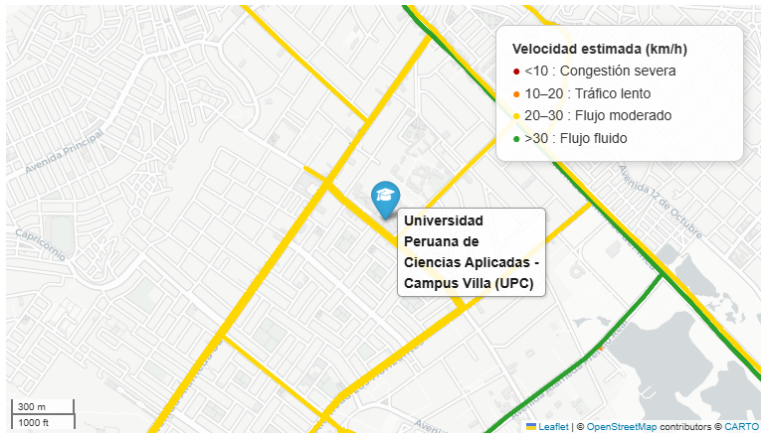


Figura: Lunes 7pm

Extensiones

Extensiones

- ▶ Infinitos tipos.
- ▶ Dinámica y aleatoridad.
- ▶ Estimación de los parámetros.
- ▶ Determinar solución interior.

Infinitos tipos

Problema de Kantorovich

Caso continuo: X, Y espacios Polacos (completos, separables).

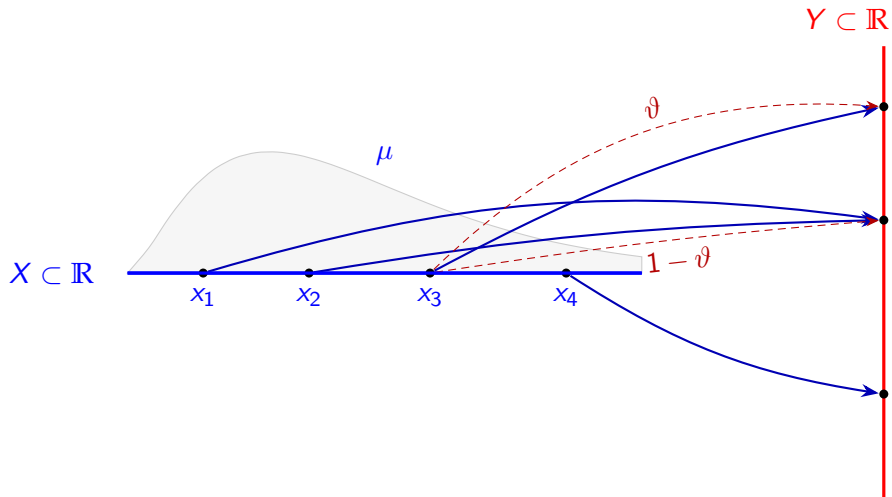
$$\inf \left\{ \int_{X \times Y} c(x, y) d\pi(x, y) : \pi \in \Pi(\mu, \nu) \right\}.$$

con

$$\Pi(\mu, \nu) = \left\{ \pi \in \Delta(X \times Y) : \begin{aligned} \pi(A \times Y) &= \mu(A), \\ \pi(X \times B) &= \nu(B), \quad \forall A, B \text{ Borelianos} \end{aligned} \right\} \quad (7)$$

y $c : X \times Y \rightarrow \overline{\mathbb{R}}_+$ semi-continua inferior.

Matching between bounded subsets of \mathbb{R}



Derivada de Radon-Nikodym

Dado un espacio medible y dos medidas π y μ , si $\pi \ll \mu$ es decir, π es absolutamente continua con respecto a μ :

$\mu(A) = 0 \implies \pi(A) = 0$, entonces existe una función f tal que:

$$\pi(A) = \int_A f \, d\mu, \quad \forall A \text{ Boreliano.}$$

A esta función f se le llama la **derivada de Radon-Nikodym** de π respecto a μ , y se denota:

$$f = \frac{d\pi}{d\mu}.$$

En otras palabras, f es la **densidad de la medida** π con respecto a μ .

Medida de conteo

Sea $X = \{x_1, \dots, x_N\}$.

- ▶ La **medida de conteo** C_X se define como:

$$C_X(A) = \#(A), \quad \forall A \subseteq X.$$

- ▶ Si μ es una medida finita sobre X , entonces se cumple:

$$\mu \ll C_X,$$

ya que $C_X(A) = 0$ si y solo si $A = \emptyset$, y por definición de medida:

$$\mu(\emptyset) = 0.$$

- ▶ Por el teorema de Radon-Nikodym, existe una densidad $f = \frac{d\mu}{dC_X}$ tal que:

$$\mu(A) = \int_A f(x) dC_X(x) = \sum_{x \in A} f(x).$$

Densidades en el caso discreto

- Sea μ una medida de probabilidad sobre $X = \{x_1, \dots, x_N\}$:

$$\mu = \sum_{i=1}^N \mu_i \delta_{x_i}, \quad \text{con } \mu_i = \left(\frac{d\mu}{dC_X} \right) (x_i).$$

O sea $f(x) = \mu_i$ si $x = x_i$, 0 caso contrario.

- Sea ν una medida de probabilidad sobre $Y = \{y_1, \dots, y_L\}$:

$$\nu = \sum_{j=1}^L \nu_j \delta_{y_j}, \quad \text{con } \nu_j = \left(\frac{d\nu}{dC_Y} \right) (y_j).$$

- Entonces, toda medida π sobre $X \times Y$ tiene la forma:

$$\pi = \sum_{i=1}^N \sum_{j=1}^L \pi_{ij} \delta_{(x_i, y_j)}, \quad \pi_{ij} = \left(\frac{d\pi}{d(C_X \otimes C_Y)} \right) (x_i, y_j).$$

Costo cuadrático: formulación discreta vs integral

El costo cuadrático heterogéneo discreto es:

$$\sum_{i=1}^N \sum_{j=1}^L c_{ij} \pi_{ij} + a_{ij} \pi_{ij}^2$$

que se corresponde con:

$$\int_{X \times Y} \left[c(x, y) \left(\frac{d\pi}{d[C_X \otimes C_Y]} \right) + a(x, y) \left(\frac{d\pi}{d[C_X \otimes C_Y]} \right)^2 \right] d[C_X \otimes C_Y].$$

Extensión al caso continuo: modelo 1

Para extender el modelo, reemplazamos C_X y C_Y por las medidas de Lebesgue \mathcal{L}_X y \mathcal{L}_Y . Suponemos que $\pi \ll \mathcal{L}_X \otimes \mathcal{L}_Y$ y definimos $\psi = \frac{d\pi}{d(\mathcal{L}_X \otimes \mathcal{L}_Y)}$. Entonces el problema se convierte en:

$$\min_{\pi \in \Pi(\mu, \nu)} \int_{X \times Y} c(x, y) d\pi(x, y) + \int_{X \times Y} a(x, y) \psi(x, y)^2 dx dy.$$

Se asume que $a(x, y) \in L^\infty(X \times Y)$ y $\psi \in L^2(X \times Y)$:

- ▶ $f \in L^\infty(X \times Y)$ si existe $M > 0$ tal que $|f(x, y)| \leq M$ casi en todas partes. Es decir, f está esencialmente acotada.
- ▶ $g \in L^2(X \times Y)$ si $\int_{X \times Y} |g(x, y)|^2 dx dy < \infty$. Es decir, g es cuadrado-integrable.

Extensión al caso continuo: modelo 2

$$\inf_{\pi \in M_+(X \times Y)} \left\{ \int_{X \times Y} \left(a(x, y) f^2(x, y) + c(x, y) f(x, y) \right) dx dy \right. \\ \left. + \|F - f_\mu\|_{L^2(X)}^2 + \|G - f_\nu\|_{L^2(Y)}^2 \right\}$$

- ▶ $M_+(X \times Y)$ son medidas positivas sobre el producto cartesiano $X \times Y$,
- ▶ f es la derivada de R.N. de π con respecto a Lebesgue $dx dy$,
- ▶ $a, c : X \times Y \rightarrow \mathbb{R}_+$ son funciones positivas,
- ▶ f_μ es la densidad respecto a Lebesgue de μ : $d\mu = f_\mu dx$, análogo con f_ν : $d\nu = f_\nu dy$,
- ▶ μ, ν medidas fijas en $M_+(X)$ e $M_+(Y)$ respectivamente,
- ▶ $\|\cdot\|_{L^2(\Omega)}$ es la norma L^2 sobre un espacio Ω ,
- ▶ $F = \int_Y f(x, y) dy$ y $G = \int_X f(x, y) dx$.

Extensión al caso continuo: modelo 2

- ▶ Derivación hecha con Carlos Cosentino (Matemáticas PUCP → Paris Dauphine).
- ▶ ¿Existencia? ¿Unicidad?
- ▶ Hipótesis sobre los agentes involucrados: funciones, espacios...

Extensión al caso continuo: modelo 2

Theorem

(Según [Brezis, 2010]) Sea E un espacio de Banach reflexivo, y sea $A \subset E$ un subconjunto cerrado y convexo. Supongamos que $\varphi : A \rightarrow (-\infty, \infty]$ es una función convexa, semicontinua inferior y no idénticamente infinita. Si se cumple alguna de las siguientes condiciones:

- ▶ A es acotado, o
- ▶ $\lim_{\substack{x \in A \\ \|x\| \rightarrow \infty}} \varphi(x) = \infty,$

entonces φ alcanza su mínimo sobre A ; es decir, existe $x^* \in A$ tal que

$$\varphi(x^*) = \inf_{x \in A} \varphi(x).$$

Definiciones clave

- ▶ **Espacio de Banach:** espacio vectorial normado completo; es decir, toda sucesión de Cauchy converge en la norma.
- ▶ **Semicontinua inferior:** una función $\varphi : E \rightarrow (-\infty, \infty]$ se dice *semicontinua inferior* (s.c.i.) en un punto $x_0 \in E$ si:

$$\liminf_{x \rightarrow x_0} \varphi(x) := \sup_{\delta > 0} (\inf \{ \varphi(x) : x \in B(x_0, \delta) \}) \geq \varphi(x_0).$$

Esto significa que, cerca de x_0 , los valores de $\varphi(x)$ no caen *abruptamente por debajo* de su valor en x_0 . No da saltos hacia abajo.

- ▶ **Espacio reflexivo:** un espacio de Banach E es reflexivo si el morfismo natural $E \hookrightarrow E^{**}$ es sobreyectivo, es decir, $E \cong E^{**}$.
- ▶ **Todo espacio de Hilbert es reflexivo** porque el isomorfismo $E \cong E^*$ inducido por el producto interno es isométrico y biyectivo, lo que implica $E \cong E^{**}$.

Extensión al caso continuo: modelo 2

- ▶ X e Y acotados, y normalizado a medida de Lebesgue 1.
- ▶ $A = L^2_+$, que es un subconjunto convexo y cerrado de L^2 es de Hilbert (y por ende reflexivo).
- ▶ Tenemos que probar que $\varphi : f \rightarrow \mathbb{E}[af^2 + cf] + \|F - f_\mu\|_{L^2(X)}^2 + \|G - f_\nu\|_{L^2(Y)}^2$ es convexa y l.s.c.
- ▶ Compactificar (acotar) A : propósitos computacionales, o acotar f .

L^2_+ convexo y cerrado

Es convexo trivialmente. Para probar que $L^2_+ \subset L^2$ es un conjunto cerrado, consideremos una sucesión $\{f_n\} \subset L^2_+$ que converge en norma L^2 hacia alguna función $f \in L^2$. Queremos demostrar que $f \in L^2_+$, es decir, que $f(x) \geq 0$ casi en todas partes. Supongamos, por contradicción, que $f \notin L^2_+$. Entonces existe un conjunto medible $S \subset X$ tal que $\mu(S) > 0$ y

$$f(x) < -\frac{1}{k} \quad \text{para todo } x \in S,$$

para algún $k \in \mathbb{N}$. Entonces, para todo n ,

$$\|f_n - f\|_{L^2}^2 \geq \int_S |f_n(x) - f(x)|^2 dx \geq \int_S \left(\frac{1}{k}\right)^2 dx = \frac{1}{k^2} \mu(S) > 0.$$

Esto contradice la hipótesis de que $f_n \rightarrow f$ en L^2 . Por tanto, la función f no puede ser negativa en un conjunto de medida positiva.

Convexidad de la aplicación φ

- $\|f\sqrt{a}\|_{L^2}^2 = \int_{X \times Y} a(x, y) f^2(x, y) dx dy$ es (estrictamente)³ convexa pues

$$\begin{aligned} & \int_{X \times Y} a(x, y) (\theta f + (1 - \theta)g)^2 dx dy \\ & < \int_{X \times Y} a(x, y) (\theta f^2 + (1 - \theta)g^2) dx dy. \end{aligned}$$

- Las aplicaciones $I : f \rightarrow F$ y $J : f \rightarrow G$ son lineales. Luego, como $\|\cdot\|$ es convexa, concluimos la convexidad de $\|I(f) - f_\mu\|_{L^2(X)}^2 + \|J(f) - f_\mu\|_{L^2(Y)}^2$.
- El término $\langle f, c \rangle = \int_{X \times Y} c(x, y) f(x, y) dx dy$ es lineal.

³Si imponemos positividad en casi todo punto.

Semi-continuidad inferior (l.s.c.)

- ▶ $\langle f, c \rangle$ es lineal.
- ▶ I y J son contracciones:

$$\|I(f)\|_{L^2(X)}^2 = \int_X \left| \int_Y f(x, y) d\mathcal{L}(y) \right|^2 d\mathcal{L}(x) \quad (8)$$

$$\leq \mathcal{L}(Y) \int_X \int_Y |f(x, y)|^2 \underbrace{d\mathcal{L}(y) d\mathcal{L}(x)}_{\doteq dx dy} \quad (9)$$

$$= \underbrace{\mathcal{L}(Y)}_{=1} \|f\|_{L^2(X \times Y)}^2. \quad (10)$$

- ▶ El término $\|f\sqrt{a}\|_{L^2}^2$ es l.s.c. por el Lema de Fatou, usando que $a(x, y) \geq 0$.

Lema de Fatou y l.s.c. de $\|f\sqrt{a}\|_{L^2}^2$

Supongamos $f_n \rightarrow f$ en $L^2(X \times Y)$. Tomando una subsucesión (que no renombramos) tenemos $f_n(x, y) \rightarrow f(x, y)$ a.e. Entonces,

$$g_n(x, y) := a(x, y) f_n^2(x, y) \geq 0, \quad g(x, y) := a(x, y) f^2(x, y).$$

Por Fatou,

$$\begin{aligned} \int_{X \times Y} g(x, y) d\mathcal{L}(x) d\mathcal{L}(y) &= \int_{X \times Y} a(x, y) f^2(x, y) d\mathcal{L}(x) d\mathcal{L}(y) \\ &\leq \int_{X \times Y} \liminf_{n \rightarrow \infty} g_n(x, y) d\mathcal{L}(x) d\mathcal{L}(y) \\ &\leq \liminf_{n \rightarrow \infty} \int_{X \times Y} a(x, y) f_n^2(x, y) d\mathcal{L}(x) d\mathcal{L}(y). \end{aligned}$$

Equivalente en norma:

$$\|f\sqrt{a}\|_{L^2(X \times Y)}^2 \leq \liminf_{n \rightarrow \infty} \|f_n\sqrt{a}\|_{L^2(X \times Y)}^2.$$

Existencia y unicidad

- ▶ Para aplicar el teorema 2, o bien imponemos a la optimización

$$\{f : |f| \leq M \text{ c.s.}\}$$

- ▶ O bien tomamos $a(x, y)$ uniformemente positiva:
 $a(x, y) > \delta > 0$ para todo $(x, y) \in X \times Y$.
- ▶ En efecto:

$$\varphi(f) > \delta \|f\|_{L^2(X \times Y)}.$$

Candidato a óptimo

- Consideremos el caso a uniformemente positiva.
- Por un argumento variacional: $\left. \frac{d}{dt} \varphi(f + tg) \right|_{t=0} \geq 0$, llegamos a

$$f(x, y)a(x, y) + I(f) + J(f) \geq f_\mu + f_\nu - \frac{c}{2} \text{ c.s.}$$

- Esto equivale a, asumiendo interioridad (similar al caso discreto)

$$f = \left[\frac{f_\mu + f_\nu - c/2}{a} - \frac{I(f) + I(g)}{a} \right] = B[f].$$

- Si consideramos únicamente $B : L^2 \rightarrow L^2$, y $a > M > 2$, como B es una contracción, por el teorema del punto fijo de Banach, existe un único f^* solución.

$$\begin{aligned} \|B(f) - B(g)\|_{L^2(X \times Y)} &\leq \frac{\|I(g - f)\|_{L^2(X)} + \|J(g - f)\|_{L^2(Y)}}{a(x, y)} \\ &\leq \frac{2}{M} \|g - f\|_{L^2(X \times Y)}. \end{aligned}$$

Estimación de los parámetros

Algunas ideas

- Nelder-mead para

$$\min_{\mathbf{c}, \mathbf{a}} \left\{ \sum_{i,j} \left(\pi_{ij}^{obs} - \underbrace{\pi_{ij}^*(\mathbf{c}, \mathbf{a})}_{A^{-1}b} \right)^2 \right\}$$

dados los demás parámetros $\epsilon, \delta, \mu, \nu, \alpha$.

- Especificaciones: $c_{ij} = \beta_0 + \beta_1 X_{1ij} + \dots + \beta_k X_{kij} + \eta_{ij} \dots$
- $\lambda a_{ij} \gg \tilde{a}_{ij}$? $\mu, \nu \rightarrow \infty$, $[c_{ij}] \simeq c \cdot I_{NL \times NL}$.
- Por ej.:

$c(i, j) = \Psi(\text{dificultad examen de admisi3n}_j, \text{rendimiento}_i, \dots$
 $\text{gusto deporte}_i, \text{infraestructura}_j) \dots$

Extensión dinámica del modelo de matching con congestión

- ▶ Consideramos una secuencia de planes de matching $\{\pi_t\}_{t=1}^T$.
- ▶ Los parámetros $c_{ij}^t, a_{ij}^t, \mu_i^t, \nu_j^t$ varían estocásticamente.
- ▶ Penalizamos cambios abruptos en los matchings mediante una norma cuadrática intertemporal:

$$\min_{\{\pi_t\}} \mathbb{E} \left[\sum_{t=1}^T \left(\sum_{i,j} c_{ij}^t \pi_{ij}^t + a_{ij}^t (\pi_{ij}^t)^2 \right. \right. \\ \left. \left. + \sum_i \epsilon_i^t \left(\sum_j \pi_{ij}^t - \mu_i^t \right)^2 + \sum_j \delta_j^t \left(\sum_i \pi_{ij}^t - \nu_j^t \right)^2 \right) \right. \\ \left. + \lambda \sum_{t=2}^T \|\pi_t - \pi_{t-1}\|_F^2 \right]$$

Dinámica estocástica de parámetros

- Supongamos que los parámetros evolucionan de forma autorregresiva:

$$c_{ij}^{t+1} = \rho_c c_{ij}^t + \eta_{ij}^{t+1},$$

$$a_{ij}^{t+1} = \rho_a a_{ij}^t + \zeta_{ij}^{t+1},$$

$$\mu_i^{t+1} = \rho_\mu \mu_i^t + \xi_i^{t+1},$$

$$\nu_j^{t+1} = \rho_\nu \nu_j^t + \vartheta_j^{t+1}$$

- Las innovaciones $\eta, \zeta, \xi, \vartheta$ son procesos de media cero con varianza constante.

Control óptimo estocástico con inversión en infraestructura

- ▶ El planificador central puede afectar los costos mediante decisiones de inversión en infraestructura:

$$\begin{aligned}c_{ij}^{t+1} &= \rho_c c_{ij}^t + \beta u_{ij}^t + \eta_{ij}^{t+1}, \\a_{ij}^{t+1} &= \rho_a a_{ij}^t + \beta' u_{ij}^t + \zeta_{ij}^{t+1}\end{aligned}$$

- ▶ u_{ij}^t es la variable de control: inversión que mejora la conectividad entre nodos i y j .
- ▶ El objetivo es maximizar la utilidad social agregada descontada:

$$\max_{\{u_t\}} \mathbb{E} \left[\sum_{t=1}^T \rho^t \left(F(\pi_t) - \gamma \|u_t\|_F^2 \right) \right]$$

$$\rho \in (0, 1).$$

Gracias

Referencias I



Abdulkadiroğlu, A. and Sönmez, T. (2003).

School Choice: A Mechanism Design Approach.

The American Economic Review, 93(3):729–747.



Alman, J., Duan, R., Williams, V. V., Xu, Y., Xu, Z., and Zhou, R. (2025).

More asymmetry yields faster matrix multiplication.

In *Proceedings of the 2025 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2005–2039, Philadelphia. SIAM.



Anaya-Montes, M. and Gravelle, H. (2024).

Health insurance system fragmentation and COVID-19 mortality: Evidence from Peru.

PLOS ONE, 19(8):e0309531.



Arrieta, A. and Guillén, J. (2017).

Output congestion leads to compromised care in peruvian public hospital neonatal units.

Health Care Management Science, 20(2):209–221.



Artstein-Avidan, S., Giannopoulos, A., and Milman, V. D. (2015).

Asymptotic Geometric Analysis, Part I, volume 202 of *Mathematical Surveys and Monographs*.

American Mathematical Society.



Bendezu-Quispe, G., Mari-Huarache, L. F., Taype-Rondan, A., Mejia, C. R., and Inga-Berrosapi, F. (2020).

Effect of rural and marginal urban health service on the physicians' perception of primary health care in peru.

Revista Peruana de Medicina Experimental y Salud Pública, 37(4):636–644.



Brezis, H. (2010).

Functional Analysis, Sobolev Spaces and Partial Differential Equations.

Springer, New York.

Referencias II



Carlier, G., Dupuy, A., Galichon, A., and Sun, Y. (2020).
SISTA: Learning Optimal Transport Costs under Sparsity Constraints.
arXiv preprint arXiv:2009.08564.
Submitted on 18 Sep 2020, last revised 21 Oct 2020.



Coppersmith, D. and Winograd, S. (1990).
Matrix multiplication via arithmetic progressions.
Journal of Symbolic Computation, 9(3):251–280.



Davie, A. M. and Stothers, A. J. (2013).
Improved bound for complexity of matrix multiplication.
Proceedings of the Royal Society of Edinburgh, Section A: Mathematics, 143(2):351–369.



Dupuy, A. and Galichon, A. (2022).
A Note on the Estimation of Job Amenities and Labor Productivity.
Quantitative Economics, 13:153–177.



Echenique, F., Immorlica, N., and Vazirani, V. V. (2023).
Online and Matching-Based Market Design.
Cambridge University Press.



Echenique, F. and Yenmez, M. B. (2015).
How to Control Controlled School Choice.
The American Economic Review, 105(8):2679–2694.



EsSalud (2025a).
Dashboard de indicadores fonafe y tablero estratégico.



EsSalud (2025b).
Tablero de diferimento de citas.

Referencias III



Gale, D. and Shapley, L. S. (1962).
College Admissions and the Stability of Marriage.
The American Mathematical Monthly, 69(1):9–15.



Galichon, A. (2016).
Optimal Transport Methods in Economics.
Princeton University Press.



Galichon, A. (2021).
The Unreasonable Effectiveness of Optimal Transport in Economics.
Preprint submitted on 12 Jan 2023.



Granot, F. and Skorin-Kapov, J. (1990).
Some proximity and sensitivity results in quadratic integer programming.
Mathematical Programming, 47(1):259–268.



Hochbaum, D. S. and Shanthikumar, J. G. (1990).
Convex separable optimization is not much harder than linear optimization.
Journal of the ACM, 37(4):843–862.



Huang, J. (2018).
Practical fast matrix multiplication algorithms.
Ph.d. dissertation, University of Texas at Austin, Austin, TX.



Hylland, A. and Zeckhauser, R. (1979).
The Efficient Allocation of Individuals to Positions.
The Journal of Political Economy, 87(2):293–314.



Izmailov, A. F. and Solodov, M. V. (2023).
Convergence rate estimates for penalty methods revisited.
Computational Optimization and Applications, 85(3):973–992.

Referencias IV



Karstadt, E. and Schwartz, O. (2017).

Matrix multiplication, a little faster.

In Proceedings of the 29th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA), pages 101–110, New York. ACM.



Kelso, A. S. and Crawford, V. P. (1982).

Job Matching, Coalition Formation, and Gross Substitutes.
Econometrica, 50(6):1483.



Koutis, I., Miller, G. L., and Peng, R. (2012).

A fast solver for a class of linear systems.

Communications of the ACM, 55(10):99–107.



Pan, V. Y. (1982).

Trilinear aggregating with implicit canceling for a new acceleration of matrix multiplication.
Computers & Mathematics with Applications, 8(1):23–34.



Park, J. and Boyd, S. (2017).

A semidefinite programming method for integer convex quadratic minimization.
Optimization Letters, 11(4):725–739.



Peng, R. and Vempala, S. S. (2024).

Solving sparse linear systems faster than matrix multiplication.
Communications of the ACM, 67(7):79–86.



Roth, A. E. (1982).

The Economics of Matching: Stability and Incentives.
Mathematics of Operations Research, 7(4):617–628.

Referencias V



Roth, A. E. and Sotomayor, M. A. O. (1990).

Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis, volume 18 of *Econometric Society Monographs*.

Cambridge University Press.



Soto, A. (2019).

Barreras para una atención eficaz en los hospitales de referencia del ministerio de salud del Perú: atendiendo pacientes en el siglo xxi con recursos del siglo xx.

Revista Peruana de Medicina Experimental y Salud Pública, 36(2):304.



Strassen, V. (1969).

Gaussian elimination is not optimal.

Numerische Mathematik, 13(4):354–356.



Strassen, V. (1986).

The asymptotic spectrum of tensors and the exponent of matrix multiplication.

In *27th Annual Symposium on Foundations of Computer Science (SFCS 1986)*, pages 49–54. IEEE.



Velasquez, A. (2020).

Consideraciones éticas del aseguramiento universal de salud en el Perú.

Technical report, Universidad Antonio Ruiz de Montoya, Lima, Perú.



Villani, C. (2009).

Optimal Transport: Old and New, volume 338 of *Grundlehren der mathematischen Wissenschaften*.

Springer.



Williams, V. V. (2012).

Multiplying matrices faster than coppersmith-winograd.

In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC)*, pages 887–898, New York. ACM.

Referencias VI



Winograd, S. (1971).

On multiplication of 2×2 matrices.

Linear Algebra and its Applications, 4(4):381–388.



World Bank (2020).

Health at a glance: Latin america and the caribbean 2020.

<https://documents1.worldbank.org/curated/en/383471608633276440/pdf/Health-at-a-Glance-Latin-America-and-the-Caribbean-2020.pdf>.

Accessed: 18 March 2025.



Zhan, S. (2005).

On the determinantal inequalities.

Journal of Inequalities in Pure and Applied Mathematics, 6(4):Article 105.