1 Lecture 1: optimal assigment

n_x workers of type x, $x \in X$
m_y firms of type y, $y \in Y$

$\sum_x n_x = \sum_y m_y$

Assume if x matches with a firm of type y, and they decide on a wage w_{xy}, then
utility of x: $\alpha_{xy} + w_{xy}$
(net) profit of y: $\gamma_{xy} - w_{xy}$

Joint surplus of xy matched together: $\Phi_{xy} = (\alpha_{xy} + w_{xy}) + \left(\gamma_{xy} - w_{xy}\right) = \alpha_{xy} + \gamma_{xy}$.

2 points of view.
=> central planner / optimality
=> decentralized economy / equilibrium

1.1 1. Optimal assignment

Decision variable $\mu_{xy} \geq 0$ = number of workers of type x assigned to firms of type y.

Conditions on μ_{xy}:
$n_x = \sum_y \mu_{xy}$
$m_y = \sum_x \mu_{xy}$

Social welfare is

$$\sum_{xy} \mu_{xy} \Phi_{xy}$$

We are able to formulate the optimal assignment problem:

$$\begin{aligned} \max_{(\mu_{xy}) \geq 0} \quad & \sum_{xy} \mu_{xy} \Phi_{xy} \\ s.t. \quad & \sum_{y} \mu_{xy} = n_x \\ & \sum_{x} \mu_{xy} = m_y \end{aligned}$$

$$\begin{aligned} \max_{z \geq 0} \quad & z^\top c \\ s.t. \quad & Az = d \end{aligned}$$

Numpy uses the row-major order to vectorize matrices. If $\mu = \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \\ \mu_{31} & \mu_{32} \end{pmatrix}$, then

$vec_R(\mu) = (\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22}, \mu_{31}, \mu_{32})$.

Now, how do we reinterpret the constraint

$$\sum_y \mu_{xy} = n_x$$

We have

$$A_1 vec_R(\mu) = n$$

where

$$A_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

written in a block-wise manner,

$$\begin{aligned} A_1 &= \begin{pmatrix} 1_Y^\top & 0_Y^\top & 0_Y^\top \\ 0_Y^\top & 1_Y^\top & 0_Y^\top \\ 0_Y^\top & 0_Y^\top & 1_Y^\top \end{pmatrix} \\ &= I_X \otimes 1_Y^\top \end{aligned}$$

Now let's consider

$$\sum_x \mu_{xy} = m_y$$

As a reminder, $vec_R(\mu) = (\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22}, \mu_{31}, \mu_{32})$.

$$\begin{aligned} A_2 &= \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \\ &= 1_X^\top \otimes I_Y \end{aligned}$$

Thus

$$\begin{aligned} A_1 &= I_X \otimes 1_Y^\top \\ A_2 &= 1_X^\top \otimes I_Y \end{aligned}$$

1.2 Equilibrium

Let's start with the optimization problem

$$\max_{(\mu_{xy}) \geq 0} \sum_{xy} \mu_{xy} \Phi_{xy} + \sum_x L\left(n_x - \sum_y \mu_{xy}\right) + \sum_y L\left(m_y - \sum_x \mu_{xy}\right)$$

$L(z) = \min_{u \in R} \{uz\}$

$$\max_{(\mu_{xy}) \geq 0} \sum_{xy} \mu_{xy} \Phi_{xy} + \sum_x \min_{u_x} \left\{ u_x \left(n_x - \sum_y \mu_{xy} \right) \right\} + \sum_y \min_{v_y} \left\{ v_y \left(m_y - \sum_x \mu_{xy} \right) \right\}$$

$$= \max_{(\mu_{xy}) \geq 0} \min_{u_x, v_y} \sum_{xy} \mu_{xy} \Phi_{xy} + \sum_x u_x \left(n_x - \sum_y \mu_{xy} \right) + \sum_y v_y \left(m_y - \sum_x \mu_{xy} \right)$$

$$= \min_{u_x, v_y} \sum_x u_x n_x + \sum_y v_y m_y + \max_{(\mu_{xy}) \geq 0} \sum_{xy} \mu_{xy} \left(\Phi_{xy} - u_x - v_y \right)$$

what is

$$\max_{(\mu_{xy}) \geq 0} \mu_{xy} \left(\Phi_{xy} - u_x - v_y \right)$$

this is

$$\begin{array}{ll} \min_{u_x, v_y} & \sum_x u_x n_x + \sum_y v_y m_y \\ s.t. & u_x + v_y \geq \Phi_{xy} \end{array}$$

If μ_{xy} is a solution to the primal problem and (u_x, v_y) to the dual problem, then for each xy,

$$\mu_{xy} \left(u_x + v_y - \Phi_{xy} \right) = 0$$

which means that

$$\mu_{xy} > 0 \implies u_x + v_y = \Phi_{xy}.$$

INTERPRETATION OF u and v.

Remember, if μ_{xy} is a solution to the primal problem and (u_x, v_y) to the dual problem, then

$$\sum_{xy} \mu_{xy} \Phi_{xy} = \sum_x u_x n_x + \sum_y v_y m_y$$

It suggests that u_x and v_y would be the payoff of x and y at equilibrium.

Claim: I can define a vector of wages w_{xy} such that if (u_x, v_y) is the solution to the dual problem, then

$u_x = \max_y \{\alpha_{xy} + w_{xy}\}$ is the indirect utility of a worker of type x

$v_y = \max_x \{\gamma_{xy} - w_{xy}\}$ is the indirect utility of a firm of type y

and if μ_{xy} is the solution to the primal problem, then

$\mu_{xy} > 0$ implies $u_x = \alpha_{xy} + w_{xy}$ (i.e., y is optimal for x), and $v_y = \gamma_{xy} - w_{xy}$ (i.e., x is optimal for y).

This means that

$u_x \geq \alpha_{xy} + w_{xy}$ for all x and y, and

$v_y \geq \gamma_{xy} - w_{xy}$ for all x and y.

That is

$$\gamma_{xy} - v_y \leq w_{xy} \leq u_x - \alpha_{xy}.$$

$u_x - \alpha_{xy} \geq \gamma_{xy} - v_y$ because $u_x + v_y \geq \alpha_{xy} + \gamma_{xy} = \Phi_{xy}$

Claim. If

$$\gamma_{xy} - v_y \leq w_{xy} \leq u_x - \alpha_{xy}.$$

then
$u_x = \max_y \{\alpha_{xy} + w_{xy}\}$ is the indirect utility of a worker of type x
$v_y = \max_x \{\gamma_{xy} - w_{xy}\}$ is the indirect utility of a firm of type y
and if μ_{xy} is the solution to the primal problem, then
$\mu_{xy} > 0$ implies $u_x = \alpha_{xy} + w_{xy}$ (i.e., y is optimal for x), and $v_y = \gamma_{xy} - w_{xy}$ (i.e., x is optimal for y).

Proof. $w_{xy} \leq u_x - \alpha_{xy}$ hence $u_x \geq \alpha_{xy} + w_{xy}$ hence $u_x \geq \max_y \{\alpha_{xy} + w_{xy}\}$.
Next, for $\mu_{xy} > 0$, $u_x - \alpha_{xy} = \gamma_{xy} - v_y$ – indeed, we had seen that

$$\mu_{xy} > 0 \implies u_x + v_y = \Phi_{xy}.$$

therefore $\mu_{xy} > 0$ implies $w_{xy} = u_x - \alpha_{xy}$, and this implies $u_x = \max_y \{\alpha_{xy} + w_{xy}\}$. As a consequence, $\mu_{xy} > 0$ implies $u_x = \alpha_{xy} + w_{xy}$. The same works on the other side.

2 Lecture 2. Semi-discrete optimal transport

Continuum of x's – inhabitants on the surface of a city. x=location of inhabitant. $n(x)$ is the density of inhabitants at x. $X = [0,1]^d$ is the set of location of inhabitants.

We still assume a finite number of y– location of facilities. $y_1, ..., y_J$. Assume that facility y_j has capacity m_j. Assume $\int n(x)\,dx = 1$ and $\sum_j m_j = 1$. Assume $\Phi(x, y_j) = -|x - y_j|^2 / 2$.

Central planner solution // decentralized solution.

Primal problem: define $\mu(y_j|x)$ as the share of inhabitants at location x that we send to facility y_j, and $\mu(x, y_j) = \mu(y_j|x)\, n(x)$, so that

$$\begin{aligned}
\max_{(\mu(x,y_j))\geq 0} \quad & -\int_x \sum_{j=1}^{J} \mu(x, y_j)\, |x - y_j|^2 / 2 \\
s.t. \quad & \sum_{j=1}^{J} \mu(x, y_j) = n(x) \\
& \int_x \mu(x, y_j)\, dx = m_j
\end{aligned}$$

and the dual is

$$\begin{aligned} \min_{\tilde{u}(x),\tilde{v}_j} \quad & \int \tilde{u}(x)\, n(x)\, dx + \sum_j \tilde{v}_j m_j \\ s.t. \quad & \tilde{u}(x) + \tilde{v}_j \geq -\frac{|x-y_j|^2}{2} \end{aligned}$$

Inhabitant at fountain x

$$\tilde{u}(x) = \max_j \left\{ -\frac{|x-y_j|^2}{2} - \tilde{v}_j \right\}$$

which allows us to reformulate the problem as

$$\min_{\tilde{v}_j} \int \max_j \left\{ -\frac{|x-y_j|^2}{2} - \tilde{v}_j \right\} n(x)\, dx + \sum_j \tilde{v}_j m_j$$

refomrulate this into

$$\begin{aligned} & \min_{\tilde{v}_j} \int \max_j \left\{ -\frac{|x|^2}{2} - \frac{|y_j|^2}{2} + x^\top y_j - \tilde{v}_j \right\} n(x)\, dx + \sum_j \tilde{v}_j m_j \\ = \quad & \min_{\tilde{v}_j} \int -\frac{|x|^2}{2} + \max_j \left\{ -\frac{|y_j|^2}{2} + x^\top y_j - \tilde{v}_j \right\} n(x)\, dx + \sum_j \tilde{v}_j m_j \end{aligned}$$

and $v_j = \tilde{v}_j + \frac{|y_j|^2}{2}$, which leads to

$$\begin{aligned} = \quad & \min_{v_j} \int \max_j \left\{ x^\top y_j - v_j \right\} n(x)\, dx + \sum_j v_j m_j \\ & - \sum_j \frac{|y_j|^2}{2} m_j - \int \frac{|x|^2}{2} n(x)\, dx \end{aligned}$$

so up to an irrelevant constnat, the problem is

$$\min_{(v_j)} F(v)$$

where

$$F(v) = \int \max_j \left\{ x^\top y_j - v_j \right\} n(x)\, dx + \sum_j v_j m_j$$

Gradient descent

$$v_j(t+1) = v_j(t) - \epsilon \frac{\partial F}{\partial v_j}(v)$$

we have

$$\begin{aligned}
\frac{\partial F}{\partial v_k}(v) &= \int -1\{x \text{ chooses } k\}\, n(x)\, dx + m_k \\
&= -\int 1\{x \text{ chooses } k\}\, n(x)\, dx + m_k \\
&= m_k - \Pr_n \{x \text{ chooses } k\} \\
&= \text{ excess supply for } k
\end{aligned}$$

3 Friday morning exercise session

$$\begin{aligned}
\max_{(\mu_{xy})\geq 0} \quad & \sum_{xy} \mu_{xy}\Phi_{xy} \\
s.t. \quad & \sum_{y} \mu_{xy} \leq n_x \\
& \sum_{x} \mu_{xy} \leq m_y
\end{aligned}$$

$$\begin{aligned}
\max_{(\mu_{xy})\geq 0} \quad & \sum_{xy} \mu_{xy}\Phi_{xy} + \sum_{x} \mu_{x0}\Phi_{x0} + \sum_{y} \mu_{0y}\Phi_{0y} \\
s.t. \quad & \sum_{y} \mu_{xy} + \mu_{x0} = n_x \\
& \sum_{x} \mu_{xy} + \mu_{0y} = m_y
\end{aligned}$$

but here we have $\Phi_{x0} = 0$ and $\Phi_{0y} = 0$ so the problem simplifies into

$$\begin{aligned}
\max_{(\mu_{xy})\geq 0} \quad & \sum_{xy} \mu_{xy}\Phi_{xy} \\
s.t. \quad & \sum_{y} \mu_{xy} \leq n_x \\
& \sum_{x} \mu_{xy} \leq m_y
\end{aligned}$$

$$\max_{(\mu_{xy})\geq 0} \sum_{xy} \mu_{xy}\Phi_{xy} + \sum_{x} l\left(n_x - \sum_{y} \mu_{xy}\right) + \sum_{y} l\left(m_y - \sum_{x} \mu_{xy}\right)$$

where

$$l(z) = 0 \text{ if } z \geq 0 \text{ and } l(z) = -\infty$$

we have

$$l(z) = \min_{u \geq 0} \{uz\}$$

$$
\begin{aligned}
& \max_{(\mu_{xy})\geq 0} \min_{u_x\geq 0, v_y\geq 0} \sum_{xy} \mu_{xy}\Phi_{xy} + \sum_x u_x \left(n_x - \sum_y \mu_{xy}\right) + \sum_y v_y \left(m_y - \sum_x \mu_{xy}\right) \\
= \quad & \min_{u_x\geq 0, v_y\geq 0} \max_{(\mu_{xy})\geq 0} \left\{ \sum_{xy} \mu_{xy}\Phi_{xy} + \sum_x u_x \left(n_x - \sum_y \mu_{xy}\right) + \sum_y v_y \left(m_y - \sum_x \mu_{xy}\right) \right\}
\end{aligned}
$$

we have the inner maximization which is

$$
\begin{aligned}
& \max_{(\mu_{xy})\geq 0} \left\{ \begin{array}{c} \sum_x u_x n_x + \sum_y v_y m_y \\ + \sum_{xy} \mu_{xy}\Phi_{xy} \\ + \sum_x u_x \left(- \sum_y \mu_{xy}\right) + \sum_y v_y \left(- \sum_x \mu_{xy}\right) \end{array} \right\} \\
= \quad & \sum_x u_x n_x + \sum_y v_y m_y + \max_{(\mu_{xy})\geq 0} \left\{ \begin{array}{c} \sum_{xy} \mu_{xy}\Phi_{xy} \\ + \sum_x u_x \left(- \sum_y \mu_{xy}\right) + \sum_y v_y \left(- \sum_x \mu_{xy}\right) \end{array} \right\}
\end{aligned}
$$

thus

$$
\min_{u_x\geq 0, v_y\geq 0} \sum_x u_x n_x + \sum_y v_y m_y + \max_{(\mu_{xy})\geq 0} \sum_{xy} \mu_{xy} \left(\Phi_{xy} - u_x - v_y\right)
$$

this yields

$$
\begin{aligned}
\min_{u_x\geq 0, v_y\geq 0} \quad & \sum_x u_x n_x + \sum_y v_y m_y \\
s.t. \quad & u_x + v_y \geq \Phi_{xy}
\end{aligned}
$$

$$
\begin{aligned}
\max_{\mu_{yz}\geq 0} \quad & \sum_{yz} \mu_{yz} U_{yz} - \sum_{xz} \mu_{xz} C_{xz} \\
s.t. \quad & \sum_y \mu_{yz} = \sum_x \mu_{xz} \\
& \sum_z \mu_{xz} = n_x \\
& \sum_z \mu_{yz} = m_y
\end{aligned}
$$

$$
\begin{aligned}
& \max_{\substack{\mu_{xz}\geq 0 \\ \mu_{yz}\geq 0}} \min_{p_z, u_x, v_y} \left\{ \begin{array}{c} \sum_{yz} \mu_{yz} U_{yz} - \sum_{xz} \mu_{xz} C_{xz} \\ + \sum_z p_z \left(\sum_x \mu_{xz} - \sum_y \mu_{yz} \right) \\ + \sum_x u_x \left(n_x - \sum_z \mu_{xz}\right) \\ + \sum_y v_y \left(m_y - \sum_z \mu_{yz}\right) \end{array} \right\} \\
= \quad & \min_{p_z, u_x, v_y} \max_{\substack{\mu_{xz}\geq 0 \\ \mu_{yz}\geq 0}} \left\{ \begin{array}{c} \sum_{yz} \mu_{yz} U_{yz} - \sum_{xz} \mu_{xz} C_{xz} \\ + \sum_z p_z \left(\sum_x \mu_{xz} - \sum_y \mu_{yz} \right) \\ + \sum_x u_x \left(n_x - \sum_z \mu_{xz}\right) \\ + \sum_y v_y \left(m_y - \sum_z \mu_{yz}\right) \end{array} \right\}
\end{aligned}
$$

$$\min_{p_z,u_x,v_y} \sum_x u_x n_x + \sum_y v_y m_y + \max_{\substack{\mu_{xz}\geq 0\\ \mu_{yz}\geq 0}} \left\{ \begin{array}{c} \sum_{yz} \mu_{yz} U_{yz} - \sum_{yz} \mu_{yz} p_z - \sum_{yz} \mu_{yz} v_y \\ -\sum_{xz} \mu_{xz} C_{xz} + \sum_{xz} \mu_{xz} p_z - \sum_{xz} \mu_{xz} u_x \end{array} \right\}$$

$$\min_{p_z,u_x,v_y} \sum_x u_x n_x + \sum_y v_y m_y + \max_{\substack{\mu_{xz}\geq 0\\ \mu_{yz}\geq 0}} \left\{ \begin{array}{c} \sum_{yz} \mu_{yz} \left(U_{yz} - p_z - v_y\right) \\ +\sum_{xz} \mu_{xz} \left(p_z - C_{xz} - u_x\right) \end{array} \right\}$$

thus the dual is

$$\begin{aligned} \min_{p_z,u_x,v_y} \quad & \sum_x u_x n_x + \sum_y v_y m_y \\ s.t. \quad & v_y \geq U_{yz} - p_z \\ & u_x \geq p_z - C_{xz} \end{aligned}$$

the constraints are equivalent with

$$\begin{aligned} v_y &\geq \max_{z\in Z} \{U_{yz} - p_z\} \\ u_x &\geq \max_{z\in Z} \{p_z - C_{xz}\} \end{aligned}$$

at an optimal solution (u, v) one has equalities and thus

$$\begin{aligned} v_y &= \max_{z\in Z} \{U_{yz} - p_z\} \\ u_x &= \max_{z\in Z} \{p_z - C_{xz}\} \end{aligned}$$

If x and y are matched, their efficient production is

$$\Phi_{xy} = \max_{z\in Z} \{U_{yz} - C_{xz}\}$$

4 Friday 10/27

4.1 One-dimensional matching (Becker's model)

Assume both distriubtions of x and y are continuous, but distributed on the real line. Let $n(x)$ be the density of x and $m(y)$ be the density of y. Assume

$$\int_{-\infty}^{+\infty} n(x)\, dx = \int_{-\infty}^{+\infty} m(y)\, dy = 1$$

Example: CEO model (Gabaix and Landier QJE and Tervio AER 2008).

x="talent" of the CEO – extra points of return on capital generated by the CEO (in percentage terms)

y=size (market cap) of the firm (in dollar terms)

Value created by the CEO (in dollars)

$$\Phi(x,y) = xy.$$

Central planner's problem

$$\begin{array}{ll} \max_{\mu} & \int \mu(x,y)\,\Phi(x,y)\,dxdy \\ s.t. & \int \mu(x,y)\,dy = n(x) \\ & \int \mu(x,y)\,dx = m(y) \end{array}$$

Dual problem

$$\begin{array}{ll} \min_{u(.),v(.)} & \int_{-\infty}^{+\infty} u(x)\,n(x)\,dx + \int_{-\infty}^{+\infty} v(y)\,m(y)\,dy \\ s.t. & u(x) + v(y) \geq \Phi(x,y) \end{array}$$

If (u, v) is a solution then

$$v(y) = \max_{x} \{\Phi(x,y) - u(x)\}$$

therefore the dual problem rewrites

$$\min_{u(.)} \int_{-\infty}^{+\infty} u(x)\,n(x)\,dx + \int_{-\infty}^{+\infty} \max_{x} \{\Phi(x,y) - u(x)\}\,m(y)\,dy.$$

Assume that u is differentiable; then by first order conditions in the firm's problem, we have that if y hires a ceo of talent x,

$$\frac{\partial \Phi(x,y)}{\partial x} = u'(x)$$

and therefore as $\Phi(x,y) = xy$

$$u'(x) = y.$$

Claim: u is a convex function. Indeed, we have

$$v(y) = \max_{x} \{\Phi(x,y) - u(x)\}$$

but also

$$\begin{aligned} u(x) &= \max_{y} \{\Phi(x,y) - v(y)\} \\ &= \max_{y} \{xy - v(y)\} \end{aligned}$$

hence $u(x)$ is convex. Therefore $u'(x)$ is increasing in x. Thus, if $T(x)$ is matched with x, then $T(x)$ is increasing in x. This is positive assortative matching.

We can now solve for the primal problem that is the distrubution of (x, y).

Indeed, let us define F_n and F_m the cumulative distribution function of x and y respectively.

$F_n(x) = \Pr(X \leq x)$ and $F_m(y) = \Pr(Y \leq y)$.

But we know that if (X, Y) are solution to the primal problem (a matched CEO-firm pair), then $Y = T(X)$ where $T(.)$ is increasing, and thus

$$\Pr(T(X) \leq y) = \Pr(Y \leq y) = F_m(y)$$

which implies

$$F_n\left(T^{-1}(y)\right) = \Pr\left(X \leq T^{-1}(y)\right) = F_m(y)$$

thereore, applying this to $y = T(x)$

$$F_n(x) = F_m(T(x))$$

and therefore

$$T(x) = F_m^{-1}(F_n(x)).$$

THis is expressing

$$F_n(x) = F_m(T(x)).$$

But remember, we had

$$u'(x) = T(x)$$

therefore

$$\begin{aligned} u(x) &= c + \int_0^x T(z)\,dz. \\ &= c + \int_0^x F_m^{-1}(F_n(z))\,dz \end{aligned}$$

Consider a CEO of talent x; this CEO creates value

$\Phi(x, T(x))$ – the CEO gets $u(x)$

Now consider a CEO who is slightly more talented, say of talent $x + dx$

the value created by this more talented CEO is

$\Phi(x + dx, T(x + dx)) = \Phi(x, T(x)) + \partial_x \Phi(x, T(x))\,dx + \partial_y \Phi(x, T(x))\,T'(x)\,dx$

– the CEO gets $u(x + dx) = u(x) + u'(x)\,dx = u(x) + \partial_x \Phi(x, T(x))\,dx$

Theorem (Becker): if the surplus function $\Phi(x, y)$ is supermodular, that is

$$\partial^2_{xy} \Phi(x, y) = \frac{\partial^2 \Phi(x, y)}{\partial x \partial y} \geq 0$$

then the solution to the primal problem has positive assortative matching, and hence

$$Y = T(X)$$

with T increasing and $T(x) = F_m^{-1}(F_n(x))$.

Sketch of a proof. Remember the first order conditions in the firm's problem

$$v(y) = \max_x \{\Phi(x,y) - u(x)\}$$

are

$$\frac{\partial \Phi(x, T(x))}{\partial x} = u'(x)$$

but by the second order conditions, we have

$$\frac{\partial^2 \Phi(x,y)}{\partial x^2} - u''(x) \leq 0$$

We take the first order condition in the firm's problem and derive it wrt x

$$\frac{\partial^2 \Phi(x, T(x))}{\partial x^2} + \frac{\partial^2 \Phi(x, T(x))}{\partial x \partial y} T'(x) = u''(x)$$

thus we have

$$T'(x) = \frac{u''(x) - \frac{\partial^2 \Phi(x,T(x))}{\partial x^2}}{\frac{\partial^2 \Phi(x,T(x))}{\partial x \partial y}} \geq 0.$$

A fiscal model of marriage.

If x is your gross income, then $N(x)$ is your net income and N is increasing and concave (single individual).

For a married couple of incomes x and y, then we create 2 "persons" making $\frac{x+y}{2}$ each.

The net income of the household is $2N\left(\frac{x+y}{2}\right)$

By concavity of N we have

$$2N\left(\frac{x+y}{2}\right) \geq N(x) + N(y).$$

We can therefore define the "fiscal surplus from marriage" as

$$\Phi(x,y) = 2N\left(\frac{x+y}{2}\right) - N(x) - N(y)$$

we have

$$\frac{\partial^2 \Phi(x,y)}{\partial x \partial y} = \frac{1}{2} N''\left(\frac{x+y}{2}\right) \leq 0$$

we have negative assortative matching: for a matched pair (X,Y)

$$F_m(Y) = 1 - F_n(X)$$

that is

$$Y = F_m^{-1}(1 - F_n(X))$$

4.2 Optimal transport with entropic regularization

Consider the problem

$$\begin{aligned}
\max_{(\mu_{xy})\geq 0} \quad & \sum_{xy} \mu_{xy}\Phi_{xy} - \sigma \sum_{xy} \mu_{xy} \ln \mu_{xy} \\
s.t. \quad & \sum_{y} \mu_{xy} = n_x \\
& \sum_{x} \mu_{xy} = m_y
\end{aligned}$$

where $\sigma > 0$ is a constant.

When σ is large, the solution tends towards $\mu_{xy} = n_x m_y$ which is the indepdent coupling.

Let's compute this problem

$$\begin{aligned}
& \max_{(\mu_{xy})\geq 0} \min_{u_x, v_y} \left\{ \begin{array}{c} \sum_{xy} \mu_{xy}\Phi_{xy} - \sigma \sum_{xy} \mu_{xy} \ln \mu_{xy} \\ + \sum_x u_x \left(n_x - \sum_y \mu_{xy} \right) \\ + \sum_y v_y \left(m_y - \sum_x \mu_{xy} \right) \end{array} \right\} \\
= & \min_{u_x, v_y} \max_{(\mu_{xy})\geq 0} \left\{ \begin{array}{c} \sum_{xy} \mu_{xy}\Phi_{xy} - \sigma \sum_{xy} \mu_{xy} \ln \mu_{xy} \\ + \sum_x u_x \left(n_x - \sum_y \mu_{xy} \right) \\ + \sum_y v_y \left(m_y - \sum_x \mu_{xy} \right) \end{array} \right\} \\
= & \min_{u_x, v_y} \sum_x u_x n_x + \sum_y v_y m_y + \sum_{xy} \max_{(\mu_{xy})\geq 0} \mu_{xy} \left(\Phi_{xy} - u_x - v_y - \sigma \ln \mu_{xy} \right)
\end{aligned}$$

What is

$$L_{xy}(u, v) = \max_{(\mu_{xy})\geq 0} \mu_{xy} \left(\Phi_{xy} - u_x - v_y - \sigma \ln \mu_{xy} \right)$$

by first order conditions, we have that the optimal μ_{xy} is obtained for

$$\Phi_{xy} - u_x - v_y - \sigma \ln \mu_{xy} - \sigma = 0$$

therefoer

$$\mu_{xy} = \exp\left(\frac{\Phi_{xy} - u_x - v_y - \sigma}{\sigma} \right)$$

and

$$L_{xy}(u, v) = \sigma \mu_{xy} = \sigma \exp\left(\frac{\Phi_{xy} - u_x - v_y - \sigma}{\sigma} \right)$$

and the dual problem obtains as

$$\min_{u_x, v_y} \sum_x u_x n_x + \sum_y v_y m_y + \sigma \sum_{xy} \exp\left(\frac{\Phi_{xy} - u_x - v_y - \sigma}{\sigma} \right)$$

FOC wrt u_x:

$$n_x = \sum_y \exp\left(\frac{\Phi_{xy} - u_x - v_y - \sigma}{\sigma}\right)$$

FOC wrt v_y:

$$m_y = \sum_x \exp\left(\frac{\Phi_{xy} - u_x - v_y - \sigma}{\sigma}\right).$$

4.3 Sinkhorn's algorithm / coordinate descent

Coordinate descent: fix v^k and obtain u^{k+1} that minimize wrt u_x holding v^k fixed. We have the FOC with respect to u_x that is

$$n_x = \sum_y \exp\left(\frac{\Phi_{xy} - u_x^{k+1} - v_y^k - \sigma}{\sigma}\right)$$

but this yields

$$\begin{aligned}
n_x &= \exp\left(\frac{-u_x^{k+1}}{\sigma}\right)\sum_y \exp\left(\frac{\Phi_{xy} - v_y^k - \sigma}{\sigma}\right) \\
\exp\left(\frac{u_x^{k+1}}{\sigma}\right) &= \frac{1}{n_x}\sum_y \exp\left(\frac{\Phi_{xy} - v_y^k - \sigma}{\sigma}\right) \\
u_x^{k+1} &= \sigma \log\left(\frac{1}{n_x}\sum_y \exp\left(\frac{\Phi_{xy} - v_y^k - \sigma}{\sigma}\right)\right)
\end{aligned}$$

Actually we can even compute $A_x^k = \exp\left(\frac{u_x^k}{\sigma}\right)$ and $B_y^k = \exp\left(\frac{v_y^k}{\sigma}\right)$ by

$$A_x^{k+1} = \frac{1}{n_x}\sum_y \frac{K_{xy}}{B_y^k}$$

where

$$K_{xy} = \exp\left(\frac{\Phi_{xy} - \sigma}{\sigma}\right)$$

and similarly

$$B_y^{k+1} = \frac{1}{m_y}\sum_x \frac{K_{xy}}{A_x^{k+1}}$$

Essentially we want to be able to fix the 2 iterations of the algorithm

$$\begin{aligned}
u_x^{k+1} &= \sigma \log\left(\frac{1}{n_x}\sum_y \exp\left(\frac{\Phi_{xy} - v_y^k - \sigma}{\sigma}\right)\right) \\
v_y^{k+1} &= \sigma \log\left(\frac{1}{m_y}\sum_x \exp\left(\frac{\Phi_{xy} - u_x^{k+1} - \sigma}{\sigma}\right)\right)
\end{aligned}$$

The "log-sum-exp" tricks.

$$\log\left(e^{A}+e^{B}\right)$$

$$\begin{aligned}\log\left(e^{A-t}+e^{B-t}\right) &= \log\left(\left(e^{A}+e^{B}\right)e^{-t}\right)\\ &= \log\left(e^{A}+e^{B}\right)+\log e^{-t}\\ &= \log\left(e^{A}+e^{B}\right)-t\end{aligned}$$

thus

$$\log\left(e^{A}+e^{B}\right)=\log\left(e^{A-t}+e^{B-t}\right)+t$$

Take $t=\max\left(A,B\right)$ and get

$$\sigma\log\left(e^{\frac{A}{\sigma}}+e^{\frac{B}{\sigma}}\right)=\sigma\log\left(e^{\frac{A-\max(A,B)}{\sigma}}+e^{\frac{B-\max(A,B)}{\sigma}}\right)+\max\left(A,B\right)$$

To apply this trick here, we do

$$\begin{aligned}a_{x}^{k+1} &= \max_{y}\left\{\Phi_{xy}-v_{y}^{k}-\sigma\right\}\\ u_{x}^{k+1} &= a_{x}^{k+1}+\sigma\log\left(\frac{1}{n_{x}}\sum_{y}\exp\left(\frac{\Phi_{xy}-v_{y}^{k}-a_{x}^{k+1}-\sigma}{\sigma}\right)\right)\\ b_{y}^{k+1} &= \max_{x}\left\{\Phi_{xy}-u_{x}^{k+1}-\sigma\right\}\\ v_{y}^{k+1} &= b_{y}^{k+1}+\sigma\log\left(\frac{1}{m_{y}}\sum_{x}\exp\left(\frac{\Phi_{xy}-u_{x}^{k+1}-\sigma-b_{y}^{k+1}}{\sigma}\right)\right)\end{aligned}$$