

# 'MATH+ECON+CODE' MASTERCLASS ON MATCHING MODELS, OPTIMAL TRANSPORT AND APPLICATIONS

Alfred Galichon (New York University)

Spring 2019

Wednesday: "Optimal transport II"

Block 9. Entropic regularization

- ▶ Entropic regularization
- ▶ The log-sum-exp trick
- ▶ The Iterated Proportional Fitting Procedure (IPFP)

- ▶ [OTME], Ch. 7.3
- ▶ Peyré, Cuturi, Computational Optimal Transport, Ch. 4.

- Consider the problem

$$\max_{\pi \in \mathcal{M}(p,q)} \sum_{ij} \pi_{ij} \Phi_{ij} - \sigma \sum_{ij} \pi_{ij} \ln \pi_{ij}$$

where  $\sigma > 0$ . The problem coincides with the optimal assignment problem when  $\sigma = 0$ . When  $\sigma \rightarrow +\infty$ , the solution to this problem approaches the independent coupling,  $\pi_{ij} = p_i q_j$ .

- Later on, we will provide microfoundations for this problem, and connect it with a number of important methods in economics (BLP, gravity model, Choo-Siow...). For now, let's just view this as an extension of the optimal transport problem.

- Let's compute the dual by the minimax approach. We have

$$\max_{\pi \geq 0} \min_{u,v} \sum_{ij} \pi_{ij} (\Phi_{ij} - u_i - v_j - \sigma \ln \pi_{ij}) + \sum_i u_i p_i + \sum_j v_j q_j$$

thus

$$\min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j + \max_{\pi \geq 0} \sum_{ij} \pi_{ij} (\Phi_{ij} - u_i - v_j - \sigma \ln \pi_{ij})$$

- By FOC in the inner problem, one has  $\Phi_{ij} - u_i - v_j - \sigma \ln \pi_{ij} - \sigma = 0$ , thus

$$\pi_{ij} = \exp \left( \frac{\Phi_{ij} - u_i - v_j - \sigma}{\sigma} \right)$$

and  $\pi_{ij} (\Phi_{ij} - u_i - v_j - \sigma \ln \pi_{ij}) = \sigma \pi_{ij}$ , thus the dual problem is

$$\min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j + \sigma \sum_{ij} \exp \left( \frac{\Phi_{ij} - u_i - v_j - \sigma}{\sigma} \right).$$

- After replacing  $v_j$  by  $v_j + \sigma$ , the dual is

$$\min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j + \sigma \sum_{ij} \exp \left( \frac{\Phi_{ij} - u_i - v_j}{\sigma} \right) - \sigma. \quad (V1)$$

- Claim: the problem is equivalent to

$$\min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j + \sigma \log \sum_{i,j} \exp \left( \frac{\Phi_{ij} - u_i - v_j}{\sigma} \right) \quad (\text{V2})$$

- Indeed, let us go back to the minimax expression

$$\min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j + \max_{\pi \geq 0} \sum_{ij} \pi_{ij} (\Phi_{ij} - u_i - v_j - \sigma \ln \pi_{ij})$$

we see that the solution  $\pi$  has automatically  $\sum_{ij} \pi_{ij} = 1$ ; thus we can incorporate the constraint into

$$\min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j + \max_{\pi \geq 0: \sum_{ij} \pi_{ij} = 1} \sum_{ij} \pi_{ij} (\Phi_{ij} - u_i - v_j - \sigma \ln \pi_{ij})$$

which yields (V2).

- Expression (V2) is interesting because, taking *any*  $\hat{\pi} \in M(p, q)$ , (V2) reexpresses as

$$\max_{u,v} \sum_{ij} \hat{\pi}_{ij} \left( \frac{\Phi_{ij} - u_i - v_j}{\sigma} \right) - \log \sum_{ij} \exp \left( \frac{\Phi_{ij} - u_i - v_j}{\sigma} \right)$$

therefore if the parameter is  $\theta = (u, v)$ , observations are  $ij$  pairs, and the likelihood of  $ij$  is

$$\pi_{ij}^{\theta} = \frac{\exp \left( \frac{\Phi_{ij} - u_i - v_j}{\sigma} \right)}{\sum_{ij} \exp \left( \frac{\Phi_{ij} - u_i - v_j}{\sigma} \right)}$$

- Hence, (V2) will coincide with the maximum likelihood in this model.

- Consider

$$\begin{aligned} \min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j \\ \text{s.t. } \sum_{i,j} \exp\left(\frac{\Phi_{ij} - u_i - v_j}{\sigma}\right) = 1 \end{aligned} \tag{V3}$$

- It is easy to see that the solutions of this problem coincide with (V2). Indeed, the Lagrange multiplier is forced to be one. In other words,

$$\begin{aligned} \min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j \\ \text{s.t. } \sigma \log \sum_{i,j} \exp\left(\frac{\Phi_{ij} - u_i - v_j}{\sigma}\right) = 0 \end{aligned} \tag{V3}$$



- Recall that when  $\sigma \rightarrow 0$ , one has

$$\sigma \log \left( e^{a/\sigma} + e^{b/\sigma} \right) \rightarrow \max(a, b)$$

- Indeed, letting  $m = \max(a, b)$ ,

$$\sigma \log \left( e^{a/\sigma} + e^{b/\sigma} \right) = m + \sigma \log \left( \exp \left( \frac{a-m}{\sigma} \right) + \exp \left( \frac{b-m}{\sigma} \right) \right), \quad (1)$$

and the argument of the logarithm lies between 1 and 2.

- This simple remark is actually a useful numerical recipe called the *log-sum-exp trick*: when  $\sigma$  is small, using (1) to compute  $\sigma \log \left( e^{a/\sigma} + e^{b/\sigma} \right)$  ensures the exponentials won't blow up.

- Back to the third expression, with  $\sigma \rightarrow 0$ , one has

$$\begin{aligned} \min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j \\ \text{s.t. } \max_{ij} (\Phi_{ij} - u_i - v_j) = 0 \end{aligned} \tag{V3}$$

- This is exactly equivalent with the classical Monge-Kantorovich expression

$$\begin{aligned} \min_{u,v} \sum_i u_i p_i + \sum_j v_j q_j \\ \text{s.t. } \Phi_{ij} - u_i - v_j \leq 0 \end{aligned} \tag{V3}$$

- ▶ We can compute  $\min F(x)$  by two methods:
  - ▶ Either by gradient descent:  $x(t+1) = x_t - \epsilon_t \nabla F(x_t)$ . (Steepest descent has  $\epsilon_t = 1/|\nabla F(x_t)|$ .)
  - ▶ Or by coordinate descent:  $x_i(t+1) = \arg \min_{x_i} F(x_i, x_{-i}(t))$ .
- ▶ Why do these methods converge? Let's provide some justification. We will decrease  $x_t$  by  $\epsilon d_t$ , where  $d_t$  is normalized by  $|d_t|_p := (\sum_{i=1}^n d_t^i)^{1/p} = 1$ . At first order, we have

$$F(x_t - \epsilon d_t) = F(x_t) - \epsilon d_t^T \nabla F(x_t) + O(\epsilon^1).$$

- ▶ We need to maximize  $d_t^T \nabla F(x_t)$  over  $|d_t|_p = 1$ .
  - ▶ For  $p = 2$ , we get  $d_t = \nabla F(x_t) / |\nabla F(x_t)|$
  - ▶ For  $p = 1$ , we get  $d_t = \text{sign}(\partial F(x_t) / \partial x^i)$  if  $|\partial F(x_t) / \partial x^i| = \max_j |\partial F(x_t) / \partial x^j|$ , 0 otherwise.

- Here, gradient descent is

$$u_i(t+1) = u_i(t) - \epsilon \frac{\partial F}{\partial u_i}(u(t), v(t)), \text{ and } v_j(t+1) = v_j(t) - \epsilon \frac{\partial F}{\partial v_j}(u(t), v(t))$$

while coordinate descent is

$$\frac{\partial F}{\partial u_i}(u_i(t+1), u_{-i}(t), v(t)) = 0, \text{ and } \frac{\partial F}{\partial v_j}(u(t), v_j(t+1), v_{-j}(t)) = 0.$$

- Gradient of objective function in V1:

$$\left( p_i - \sum_j \exp\left(\frac{\Phi_{ij} - u_i - v_j}{\sigma}\right), q_j - \sum_i \exp\left(\frac{\Phi_{ij} - u_i - v_j}{\sigma}\right) \right)$$

- Gradient of objective function in V2:

$$\left( p_i - \frac{\sum_j \exp\left(\frac{\Phi_{ij} - u_i - v_j}{\sigma}\right)}{\sum_{ij} \exp\left(\frac{\Phi_{ij} - u_i - v_j}{\sigma}\right)}, q_j - \frac{\sum_i \exp\left(\frac{\Phi_{ij} - u_i - v_j}{\sigma}\right)}{\sum_{ij} \exp\left(\frac{\Phi_{ij} - u_i - v_j}{\sigma}\right)} \right)$$

- Coordinate descent on objective function in V1:

$$p_i = \sum_j \exp \left( \frac{\Phi_{ij} - u_i(t+1) - v_j(t)}{\sigma} \right), q_j = \sum_i \exp \left( \frac{\Phi_{ij} - u_i(t) - v_j(t+1)}{\sigma} \right)$$

that is

$$\begin{cases} u_i(t+1) = \sigma \log \left( \frac{1}{p_i} \sum_j \exp \left( \frac{\Phi_{ij} - v_j(t)}{\sigma} \right) \right) \\ v_j(t+1) = \sigma \log \left( \frac{1}{q_j} \sum_i \exp \left( \frac{\Phi_{ij} - u_i(t)}{\sigma} \right) \right) \end{cases}$$

this is called the Iterated Fitting Proportional Procedure (IPFP), or Sinkhorn's algorithm.

- Coordinate descent on objective function in V2 does not yield a closed-form expression.

- Letting  $a_i = \exp(-u_i/\sigma)$  and  $b_j = \exp(-v_j/\sigma)$  and  $K_{ij} = \exp(\Phi_{ij}/\sigma)$ , one has  $\pi_{ij} = a_i b_j K_{ij}$ , and the procedure reexpresses as

$$\begin{cases} a_i(t+1) = p_i / (Kb(t))_i \text{ and} \\ b_j(t+1) = q_j / (K^T a(t))_j. \end{cases}$$

- The previous program is extremely fast, partly due to the fact that it involves linear algebra operations. However, it breaks down when  $\sigma$  is small; this is best seen taking a log transform and returning to  $u^k = -\sigma \log a^k$  and  $v^k = -\sigma \log b^k$ , that is

$$\begin{cases} u_i^k = \mu_i + \sigma \log \sum_j \exp \left( \frac{\Phi_{ij} - v_j^{k-1}}{\sigma} \right) \\ v_j^k = \zeta_j + \sigma \log \sum_i \exp \left( \frac{\Phi_{ij} - u_i^k}{\sigma} \right) \end{cases}$$

where  $\mu_i = -\sigma \log p_i$  and  $\zeta_j = -\sigma \log q_j$ .

- One sees what may go wrong: if  $\Phi_{ij} - v_j^{k-1}$  is positive in the exponential in the first sum, then the exponential blows up due to the small  $\sigma$  at the denominator. However, the “log-sum-exp trick” can be used in order to avoid this issue.



- Consider

$$\begin{cases} \tilde{v}_i^k = \max_j \{ \Phi_{ij} - v_j^k \} \\ \tilde{u}_j^k = \max_i \{ \Phi_{ij} - u_i^k \} \end{cases}$$

(the indexing is not a typo:  $\tilde{v}$  is indexed by  $i$  and  $\tilde{u}$  by  $j$ ).

- One has

$$\begin{cases} u_i^k = \mu_i + \tilde{v}_i^{k-1} + \sigma \log \sum_j \exp \left( \frac{\Phi_{ij} - v_j^{k-1} - \tilde{v}_i^k}{\sigma} \right) \\ v_j^k = \zeta_j + \tilde{u}_j^k + \sigma \log \sum_i \exp \left( \frac{\Phi_{ij} - u_i^k - \tilde{u}_j^k}{\sigma} \right) \end{cases}$$

and now the arguments of the exponentials are always nonpositive, ensuring the exponentials don't blow up.