Sparse Three-parameter Restricted IBP for Understanding International Trade

Melanie F. Pradier, Viktor Stojkoski, Zoran Utkovski, Ljupco Kocarev, and Fernando Perez-Cruz

December 9, 2016

Sparse Three-parameter Restricted IBP for Understanding International Trade

Melanie F. Pradier, Viktor Stojkoski, Zoran Utkovski,
Ljupco Kocarev, and Fernando Perez-Cruz

December 9, 2016

- High-dimensional count data.

Sparse Three-parameter Restricted IBP for Understanding International Trade

Melanie F. Pradier, Viktor Stojkoski, Zoran Utkovski,
Ljupco Kocarev, and Fernando Perez-Cruz

December 9, 2016

- High-dimensional count data.
- Focus on **Data Exploration**.

Motivation: Wealth of Nations

What makes some countries wealthier than others?

Motivation: Wealth of Nations

What makes some countries wealthier than others?

Classical view

- Division of labor (Ricardo, 1817; A. Smith, 1976)

Motivation: Wealth of Nations

What makes some countries wealthier than others?

Classical view

- Division of labor (Ricardo, 1817; A. Smith, 1976)
- Specialization leads to economic efficiency

Motivation: Wealth of Nations

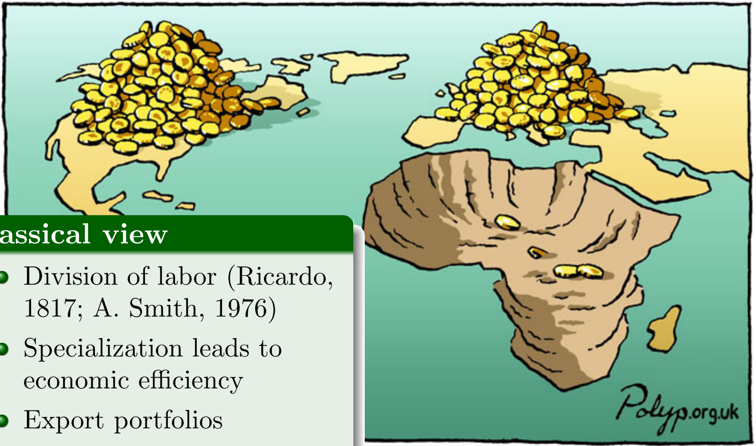
What makes some countries wealthier than others?

Classical view

- Division of labor (Ricardo, 1817; A. Smith, 1976)
- Specialization leads to economic efficiency
- Export portfolios

Motivation: Wealth of Nations

What makes some countries wealthier than others?



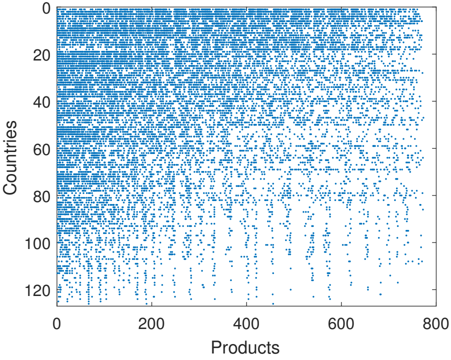
Classical view

- Division of labor (Ricardo, 1817; A. Smith, 1976)
- Specialization leads to economic efficiency
- Export portfolios

$\rightarrow$ block-structure

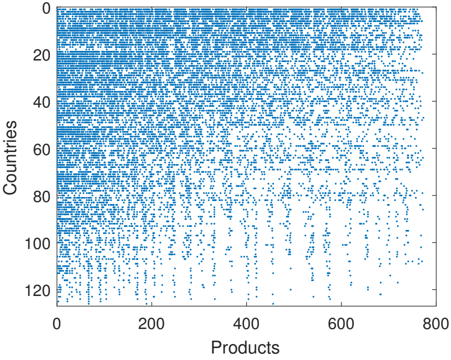
Motivation: Wealth of Nations

The reality:



Motivation: Wealth of Nations

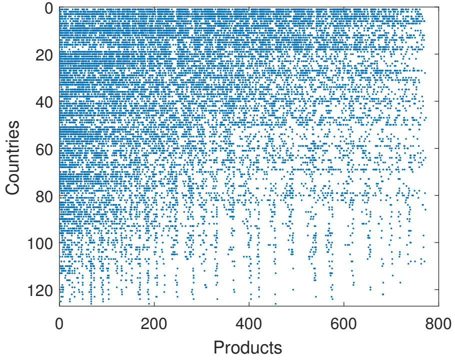
The reality:



Properties:

Motivation: Wealth of Nations

The reality:

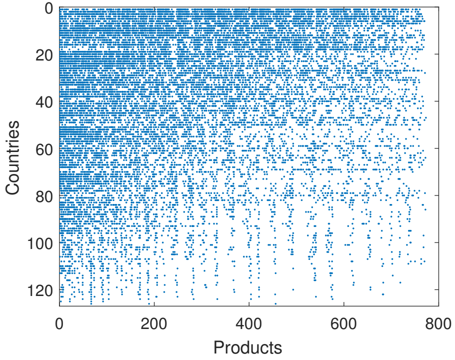


Properties:

1. Triangularity

Motivation: Wealth of Nations

The reality:

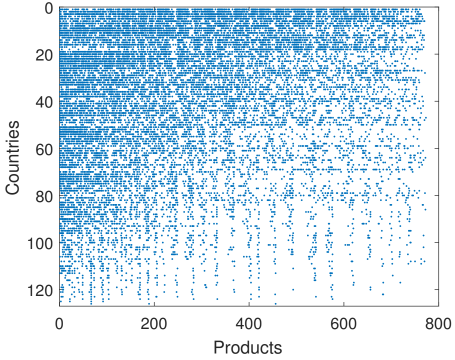


Properties:

1. Triangularity
2. $D \gg N$

Motivation: Wealth of Nations

The reality:



Properties:

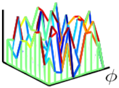
1. Triangularity
2. $D \gg N$

Our Contribution

Develop an Infinite Poisson-Gamma Model

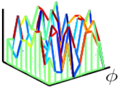
- Flexible prior
- Feature sparsity

Indian Buffet Process (Ghahramani et.al, 2006)



$$\xrightarrow[Z \sim \mathrm{IBP}(\alpha)]{} \quad Z = \begin{bmatrix} 1 & 0 & 1 & \\ 0 & 1 & 0 & \vdots \\ 1 & 1 & 0 & \\ 1 & 0 & 1 & \end{bmatrix}$$

Indian Buffet Process (Ghahramani et.al, 2006)



$$\xrightarrow[Z \sim \mathrm{IBP}(\alpha)]{} \quad Z = \begin{bmatrix} 1 & 0 & 1 & \\ 0 & 1 & 0 & \vdots \\ 1 & 1 & 0 & \\ 1 & 0 & 1 & \end{bmatrix}$$

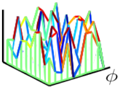
De Finetti's Representation:
(Thibaux et.al, 2007)

$$\mathbf{Z}_{n\cdot} \sim \mathrm{BeP}(\mu) \quad (1)$$

$$\mu \sim \mathrm{BP}(1, \alpha, H) \quad (2)$$

where $\mu = \sum_k \pi_k \delta_{\theta_k}$

Indian Buffet Process (Ghahramani et.al, 2006)



$$\xrightarrow{\;Z \sim \mathrm{IBP}(\alpha)\;} \qquad Z = \begin{bmatrix} 1 & 0 & 1 & \\ 0 & 1 & 0 & \vdots \\ 1 & 1 & 0 & \\ 1 & 0 & 1 & \end{bmatrix}$$

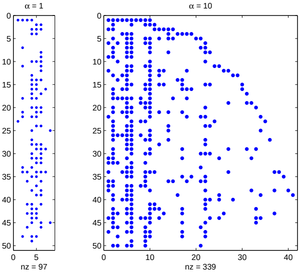
De Finetti's Representation: (Thibaux et.al, 2007)

$$\mathbf{Z}_{n\cdot} \sim \mathrm{BeP}(\mu) \qquad (1)$$

$$\mu \sim \mathrm{BP}(1, \alpha, H) \qquad (2)$$

where $\mu = \sum_k \pi_k \delta_{\theta_k}$

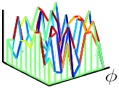
- **Disadvantage**: Mass parameter α couples both J_n and K^+



$$J_n \sim \mathrm{Poisson}(\alpha) \qquad (3)$$

$$K^+ \sim \mathrm{Poisson}\left(\alpha \sum_{n=1}^{N} \left(\frac{1}{n}\right)\right) \qquad (4)$$

Indian Buffet Process (Ghahramani et.al, 2006)



$\xrightarrow[Z \sim \mathrm{IBP}(\alpha)]{}$ $Z = \begin{bmatrix} 1 & 0 & 1 & \\ 0 & 1 & 0 & \vdots \\ 1 & 1 & 0 & \\ 1 & 0 & 1 & \end{bmatrix}$

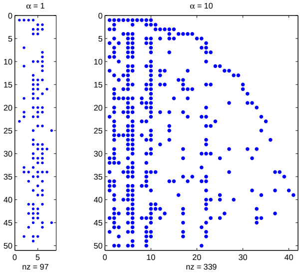
De Finetti's Representation: (Thibaux et.al, 2007)

$$\mathbf{Z}_{n\cdot} \sim \mathrm{BeP}(\mu) \qquad (1)$$

$$\mu \sim \mathrm{BP}(1, \alpha, H) \qquad (2)$$

where $\mu = \sum_k \pi_k \delta_{\theta_k}$

- **Disadvantage**: Mass parameter α couples both J_n and K^+



$$J_n \sim \mathrm{Poisson}(\alpha) \qquad (3)$$

$$K^+ \sim \mathrm{Poisson}\Big(\alpha \sum_{n=1}^{N} \Big(\frac{1}{n}\Big)\Big) \qquad (4)$$

Beyond the standard IBP

Three-parameter IBP (Teh et.al, 2007)

- More flexible distribution for stick weights

$$\mathbf{Z}_{n\cdot} \sim \mathrm{BeP}(\mu) \qquad (5)$$
$$\mu \sim \mathrm{SBP}(1, \alpha, H, c, \sigma) \qquad (6)$$

$$p\left(J_{new}\right) \sim \mathrm{Poisson}\left(\boldsymbol{\alpha}\frac{\Gamma(1+\mathbf{c})\Gamma(n+\mathbf{c}+\boldsymbol{\sigma}-1)}{\Gamma(n+\mathbf{c})\Gamma(\mathbf{c}+\boldsymbol{\sigma})}\right)$$

Beyond the standard IBP

Three-parameter IBP (Teh et.al, 2007)

- More flexible distribution for stick weights

$$\mathbf{Z}_{n\cdot} \sim \mathrm{BeP}(\mu) \quad (5)$$

$$\mu \sim \mathrm{SBP}(1, \alpha, H, c, \sigma) \quad (6)$$

$$p\left(J_{new}\right) \sim \text{Poisson}\left(\boldsymbol{\alpha}\frac{\Gamma(1+\mathbf{c})\Gamma(n+\mathbf{c}+\boldsymbol{\sigma}-1)}{\Gamma(n+\mathbf{c})\Gamma(\mathbf{c}+\boldsymbol{\sigma})}\right)$$

Beyond the standard IBP

Three-parameter IBP (Teh et.al, 2007)

- More flexible distribution for stick weights

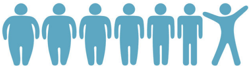
$$\mathbf{Z}_{n\cdot} \sim \text{BeP}(\mu) \quad (5)$$
$$\mu \sim \text{SBP}(1, \alpha, H, c, \sigma) \quad (6)$$

$$p\left(J_{new}\right) \sim \text{Poisson}\left(\boldsymbol{\alpha}\frac{\Gamma(1+\mathbf{c})\Gamma(n+\mathbf{c}+\boldsymbol{\sigma}-1)}{\Gamma(n+\mathbf{c})\Gamma(\mathbf{c}+\boldsymbol{\sigma})}\right)$$

Restricted IBP (Doshi-Velez et.al, 2015)

- Arbitrary prior f over J_n

$$\mathbf{Z}_{n\cdot} \sim \text{R-BeP}(\mu, f) \quad (7)$$
$$\mu \sim \text{BP}(1, \alpha, H) \quad (8)$$



Beyond the standard IBP

Three-parameter IBP (Teh et.al, 2007)

- More flexible distribution for stick weights

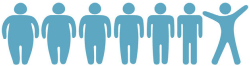
$$\mathbf{Z}_{n\cdot} \sim \text{BeP}(\mu) \quad (5)$$
$$\mu \sim \text{SBP}(1, \alpha, H, c, \sigma) \quad (6)$$

$$p\left(J_{new}\right) \sim \text{Poisson}\left(\boldsymbol{\alpha}\frac{\Gamma(1+\mathbf{c})\Gamma(n+\mathbf{c}+\boldsymbol{\sigma}-1)}{\Gamma(n+\mathbf{c})\Gamma(\mathbf{c}+\boldsymbol{\sigma})}\right)$$

Restricted IBP (Doshi-Velez et.al, 2015)

- Arbitrary prior f over J_n

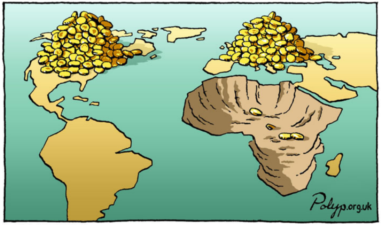
$$\mathbf{Z}_{n\cdot} \sim \text{R-BeP}(\mu, f) \quad (7)$$
$$\mu \sim \text{BP}(1, \alpha, H) \quad (8)$$

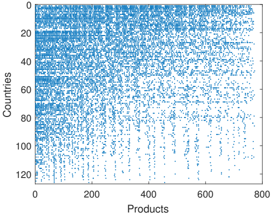


- Combination of both
- Flexible prior

Our Approach

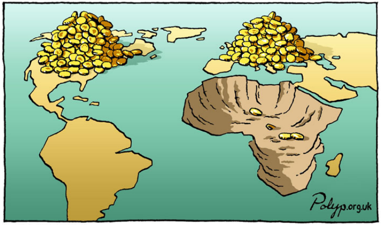
Sparse Three-Parameter Restricted Indian Buffet Process

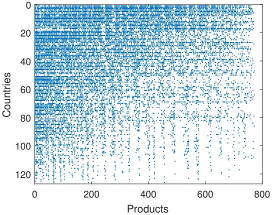




Our Approach

Sparse Three-Parameter Restricted Indian Buffet Process





Generative Model

- Let $\mathbf{X} \in \mathbb{N}^{N \times D}$

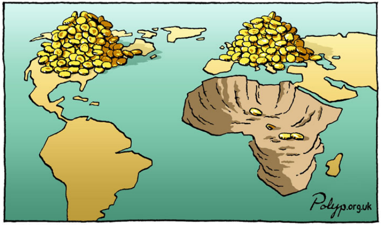
$$x_{nd} \sim \text{Poisson}(\mathbf{Z}_{n\cdot}\mathbf{B}_{\cdot d}) \quad (9)$$

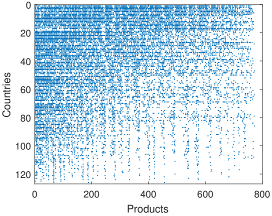
$$B_{kd} \sim \text{Gamma}\left(\alpha_B, \frac{\mu_B}{\alpha_B}\right) \quad (10)$$

$$\mathbf{Z}_{n\cdot} \sim \text{BeP}(\mu) \quad (11)$$

$$\mu \sim \text{BP}(1, \alpha, H) \quad (12)$$

Our Approach
Sparse Three-Parameter Restricted Indian Buffet Process





Generative Model

- Let $\mathbf{X} \in \mathbb{N}^{N \times D}$

$$x_{nd} \sim \text{Poisson}\big(\mathbf{Z}_{n\cdot}\mathbf{B}_{\cdot d}\big) \quad (9)$$

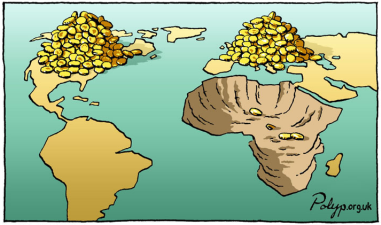
$$B_{kd} \sim \text{Gamma}\big(\alpha_B, \frac{\mu_B}{\alpha_B}\big) \quad (10)$$

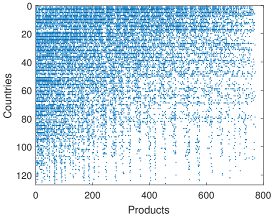
$$\mathbf{Z}_{n\cdot} \sim \text{BeP}(\mu) \quad (11)$$

$$\mu \sim \text{SBP}(1, \alpha, H, c, \sigma) \quad (12)$$

Our Approach

Sparse Three-Parameter Restricted Indian Buffet Process





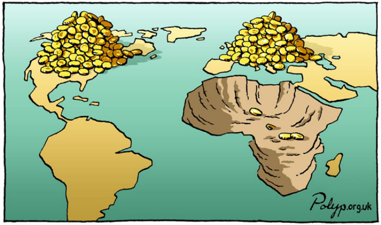
Generative Model

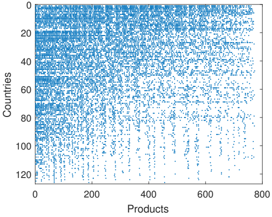
- Let $\mathbf{X} \in \mathbb{N}^{N \times D}$

$$
\begin{aligned}
x_{nd} &\sim \text{Poisson}\big(\mathbf{Z}_{n\cdot}\mathbf{B}_{\cdot d}\big) && (9)\\
B_{kd} &\sim \text{Gamma}\big(\alpha_B, \frac{\mu_B}{\alpha_B}\big) && (10)\\
\mathbf{Z}_{n\cdot} &\sim \text{R-BeP}(\mu, f) && (11)\\
\mu &\sim \text{SBP}(1, \alpha, H, c, \sigma) && (12)
\end{aligned}
$$

Our Approach

Sparse Three-Parameter Restricted Indian Buffet Process





Generative Model

- Let $\mathbf{X} \in \mathbb{N}^{N \times D}$

$$x_{nd} \sim \text{Poisson}(\mathbf{Z}_{n\cdot}\mathbf{B}_{\cdot d}) \quad (9)$$

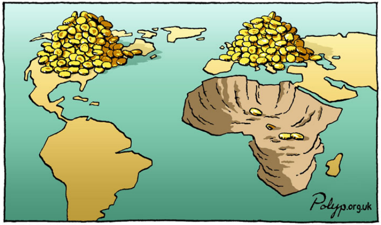
$$B_{kd} \sim \text{Gamma}\left(\alpha_B, \frac{\mu_B}{\alpha_B}\right) \quad (10)$$

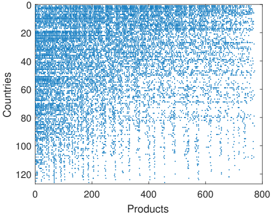
$$\mathbf{Z}_{n\cdot} \sim \text{R-BeP}(\mu, f) \quad (11)$$

$$\mu \sim \text{SBP}(1, \alpha, H, c, \sigma) \quad (12)$$

Our Approach

Sparse Three-Parameter Restricted Indian Buffet Process





Generative Model

- Let $\mathbf{X} \in \mathbb{N}^{N \times D}$

$$x_{nd} \sim \text{Poisson}(\mathbf{Z}_{n\cdot}\mathbf{B}_{\cdot d}) \quad (9)$$

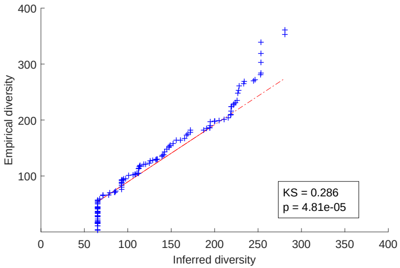
$$B_{kd} \sim \text{Gamma}\left(\alpha_B, \frac{\mu_B}{\alpha_B}\right) \quad (10)$$

$$\mathbf{Z}_{n\cdot} \sim \text{R-BeP}(\mu, f) \quad (11)$$

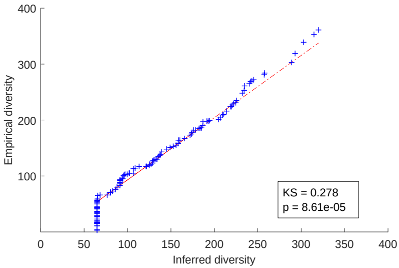
$$\mu \sim \text{SBP}(1, \alpha, H, c, \sigma) \quad (12)$$

1. 3P-IBP $\rightarrow$ high-tech features
2. R-IBP $\rightarrow$ countries inequalities
3. Sparse features ($\alpha_B < 1$) $\rightarrow$ interpretability

Results
Capturing Input Sparsity Structure



(a) S-IBP



(b) Our model S3R-IBP

Results
Interpretability

Id	$\bar{m}_k$	**Top-5 products with sorted highest weights (B_{kd})**
F1	18.27	Miscellaneous Animal Oils (0.78), Bovine and Equine Entrails (0.72), Bovine meat (0.68), Preserved Milk (0.63), Equine (0.62)
F3	14.87	Parts of Metalworking Machine Tools (0.74), Interchangeable Tool Parts (0.72), Polishing Stones (0.69), Tool Holders (0.66), Miscellaneous Metalworking Machine-Tools (0.54)
F5	11.04	Synthetic Rubber (0.87), Acrylic Polymers (0.85), Silicones (0.76), Miscellaneous Polymerization Products (0.71), Tinned Sheets (0.65)
F7	31.14	Vehicles Parts and Accessories (0.59), Cars (0.58), Iron Wire (0.53), Trucks and Vans (0.53), Air Pumps and Compressors (0.50)
...	...	...

Results
Interpretability

Top Products (decay 30%)	B_{kd}
Bovine	0.49
Miscellaneous Refrigeration Equipment	0.43
Radioactive Chemicals	0.41
Blocks of Iron and Steel	0.41
Rape Seeds	0.40
Animal meat, misc	0.39
Refined Sugars	0.38
Miscellaneous Tire Parts	0.38
Leather Accessories	0.38
Liquor	0.38
Bovine meat	0.38
Embroidery	0.37
Unmilled Barley	0.37
Dried Vegetables	0.36
Textile Fabrics Clothing Accessories	0.36
Horse Meat	0.35
Iron Bars and Rods	0.35
Analog Navigation Devices	0.35

(a) SVD

Top Products (decay 30%)	B_{kd}
Miscellaneous Animal Oils	0.78
Bovine and Equine Entrails	0.72
Bovine meat	0.68
Preserved Milk	0.63
Equine	0.62
Butter	0.58
Misc. Animal Origin Materials	0.57
Glues	0.56

(b) S3R-IBP

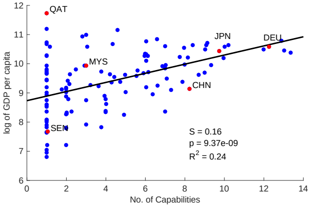
Results
Interpretability

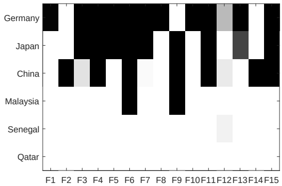
Countries in Capability Space

Results
Interpretability

Countries in Capability Space

- France = Belgium + ?
- Germany - ? = Austria
- Malaysia (Electronics) + ? → Phillipines
- Phillipines + ? → Indonesia, Vietnam

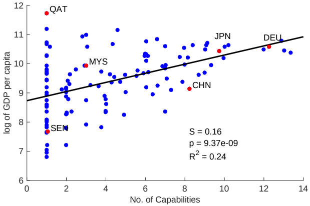


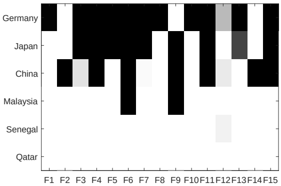


Results
Interpretability

Countries in Capability Space

- France = Belgium + Industrial Machinery
- Germany - Chemical = Austria
- Malaysia (Electronics) + Clothing → Phillipines
- Phillipines + Basic Processing → Indonesia, Vietnam





Conclusion

1. BNP model for data exploration in high-dim count data.
2. **interpretable** and **structured** solutions.
3. Analysis of productive structure of world economies.

Conclusion

1. BNP model for data exploration in high-dim count data.
2. **interpretable** and **structured** solutions.
3. Analysis of productive structure of world economies.

Future works

- **Time-dependent extension** with Markovian activation of features and smooth variation of capabilities.

Conclusion

1. BNP model for data exploration in high-dim count data.
2. **interpretable** and **structured** solutions.
3. Analysis of productive structure of world economies.

Future works

- **Time-dependent extension** with Markovian activation of features and smooth variation of capabilities.

Thank you for listening! Any question?