GLM Predictor Matrices

Analyses Europe6 and Europe7

Cecilia Valenzuela

12 December, 2020

Description

Generation of the predictor matrices for the GLM model from different sources of information:

- Eurostats, European data and statistics, transport data:
  https://ec.europa.eu/eurostat/web/transport/data/database
- United Nations Population Division, World Population Prospects 2019: R library `wpp2019`, `data(pop)`
- Distance

GLM Model

We define the migration rates m_{ij}^{e} from deme i to deme j in epoch e as a log linear function of a set of predictors $\boldsymbol{X}$, log transformed and scaled, such that

$$m_{ij}^{e} = c \exp(\beta_{pas}\delta_{pas}x_{ij.pas}^{e} + \beta_{pop}\delta_{pop}x_{ij.pop} + \beta_{dist}\delta_{dist}x_{ij.dist})$$

where β are the coefficients for the predictors that can be between $-\infty$ and ∞ (prior $\mathcal{N}(0, 4)$) and $\boldsymbol{\delta}$ are the indicators that can be 0 or 1 and denote if a predictor contributes at all.

Daily flight passengers

Passengers carried data from Eurostats. Reported *departures* values are taken for European countries (recommendation from Eurostats to avoid duplicates). For China, arrivals and departures values are considered.

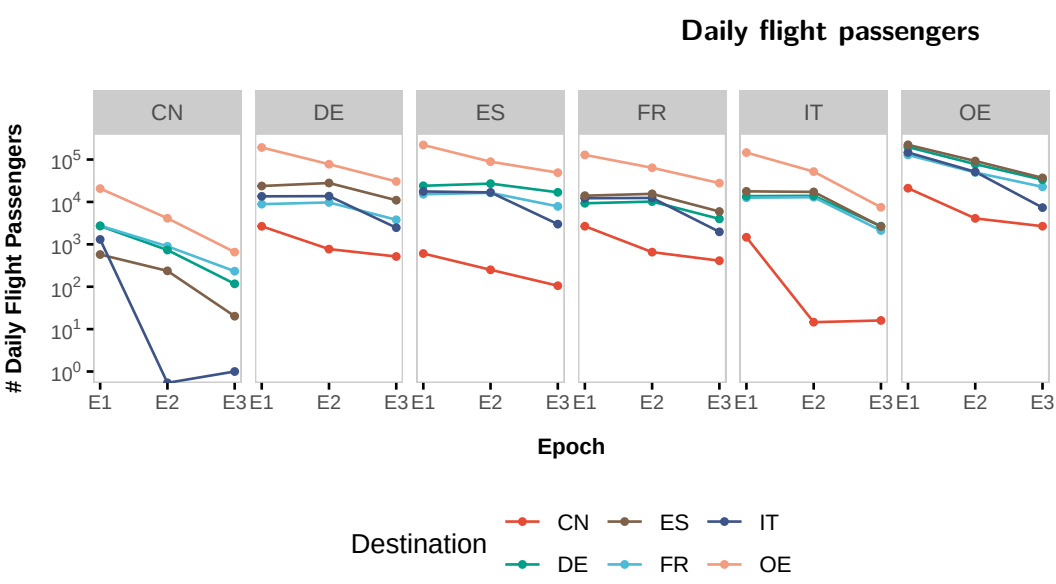
- **Epoch 1:** December 2019 - January 2020 (Wuhan lockdown 23 Jan)
- **Epoch 2:** February 2020
- **Epoch 3:** March 2020

Table 1: Flight passengers between demes, epochs 1-3

	CN	FR	DE	IT	OE	ES
CN	NA	-0.30	-0.31	-0.60	0.51	-0.93
FR	-0.31	NA	0.19	0.30	1.25	0.36
DE	-0.31	0.17	NA	0.34	1.42	0.57
IT	-0.56	0.31	0.35	NA	1.30	0.45
OE	0.52	1.25	1.43	1.31	NA	1.48
ES	-0.91	0.39	0.58	0.45	1.47	NA

	CN	FR	DE	IT	OE	ES
CN	NA	-0.75	-0.84	-3.50	-0.14	-1.29
FR	-0.88	NA	0.23	0.31	0.97	0.40
DE	-0.81	0.21	NA	0.35	1.05	0.64
IT	-2.39	0.32	0.35	NA	0.89	0.44
OE	-0.14	0.87	1.05	0.88	NA	1.12
ES	-1.27	0.42	0.62	0.43	1.10	NA

	CN	FR	DE	IT	OE	ES
CN	NA	-1.30	-1.58	-3.22	-0.88	-2.27
FR	-1.07	NA	-0.15	-0.44	0.63	0.01
DE	-0.98	-0.17	NA	-0.35	0.67	0.26
IT	-2.36	-0.41	-0.31	NA	0.10	-0.32
OE	-0.31	0.55	0.71	0.10	NA	0.75
ES	-1.62	0.12	0.43	-0.27	0.86	NA


Daily flight passengers
CN
DE
ES
FR
IT
OE
Daily Flight Passengers
10^0
10^1
10^2
10^3
10^4
10^5
E1
E2
E3
Epoch
Destination
CN
DE
ES
FR
IT
OE

Population counts

Population counts from World Population Prospects 2019, United Nations Population Division, `wpp2019` R library: **China** - 1.4393238×10^9, **France** - 6.5273512×10^7, **Germany** - 8.3783945×10^7, **Italy** - 6.0461828×10^7, **Other European** - 4.9136198×10^8 and **Spain** - 4.6754783×10^7.

A predictor matrix for source population and a predictor matrix for destination population are generated, log transformed and standardized. Constant across demes.

Table 2: Population counts source (left) and destination (right)

	CN	FR	DE	IT	OE	ES
CN	NA	1.75	1.75	1.75	1.75	1.75
FR	-0.64	NA	-0.64	-0.64	-0.64	-0.64
DE	-0.44	-0.44	NA	-0.44	-0.44	-0.44
IT	-0.70	-0.70	-0.70	NA	-0.70	-0.70
OE	0.92	0.92	0.92	0.92	NA	0.92
ES	-0.89	-0.89	-0.89	-0.89	-0.89	NA

	CN	FR	DE	IT	OE	ES
CN	NA	-0.64	-0.44	-0.7	0.92	-0.89
FR	1.75	NA	-0.44	-0.7	0.92	-0.89
DE	1.75	-0.64	NA	-0.7	0.92	-0.89
IT	1.75	-0.64	-0.44	NA	0.92	-0.89
OE	1.75	-0.64	-0.44	-0.7	NA	-0.89
ES	1.75	-0.64	-0.44	-0.7	0.92	NA

Other European population count is computed as the difference between the Europe population and the sum of the populations from France, Germany, Italy and Spain.

Distance between countries

Options: Great circle distance between centroids of each country - Which distance for other european deme? Could be the average Average pairwaise distance between airports - considers that only flight transportation matters, more work Distance taking into account population distribution inside the country. Minimum distance.