Energy System Modelling
Summer Semester 2020, Lecture 14

Dr. Tom Brown, tom.brown@kit.edu, `https://nworbmot.org/`
Karlsruhe Institute of Technology (KIT), Institute for Automation and Applied Informatics (IAI)

Table of Contents

Idea of Principal Component Analysis (PCA)

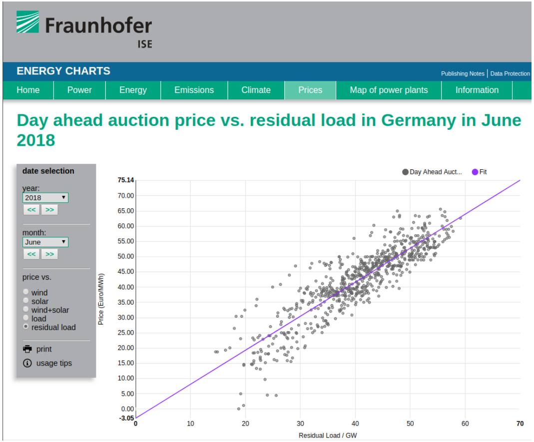
The idea of Principal Component Analysis (PCA)

Suppose we have a set of time series $x_i(t)$ for $i = 1, \ldots N$ whose means $\langle \cdot \rangle$ are centred at the origin $\langle x_i(t) \rangle = 0$ for all i.

Principal Component Analysis (PCA) is a tool to find the directions in the N-dimensional x_i space which cause the biggest variance.

We change to a new (orthonormal) basis ρ_i^k ($k = 1, \ldots N$) in N-dimensional x_i space where the first basis vector ρ^1 is in the direction of highest variance, the second ρ^2 in the next highest, etc.

We can then use this for **dimensional reduction** and ignore directions with low variance.



Source: https://energy-charts.de/

Procedure 1/2

- Calculate the **covariance matrix**:

$$\Sigma_{ij} = \langle x_i(t) x_j(t) \rangle - \langle x_i(t) \rangle \langle x_j(t) \rangle = \langle x_i(t) x_j(t) \rangle$$

  remembering that we've arranged $\langle x_i(t) \rangle = 0$. NB: The covariance matrix is **symmetric** ($\Rightarrow N$ orthogonal eigenvectors) and **positive semi-definite** ($\Rightarrow$ eigenvalues $\lambda_k \geq 0$). The diagonal entries Σ_{ii} give the variance of each $x_i(t)$.

- Find the **eigenvectors** ρ_i^k and **eigenvalues** λ_k for $k = 1, \ldots N$ of the **normalised covariance matrix** $\frac{\Sigma}{\mathrm{tr}(\Sigma)}$

$$\frac{1}{\mathrm{tr}(\Sigma)} \sum_j \Sigma_{ij} \rho_j^k = \lambda_k \rho_i^k$$

  The normalization is chosen such that $\sum_{k=1}^{N} \lambda_k = \mathrm{tr}\left(\frac{\Sigma}{\mathrm{tr}(\Sigma)}\right) = 1$.

Procedure 2/2

- Order the eigenvectors ρ_i^k and eigenvalues λ_k from highest λ_k to lowest. The value λ_k represents the share of the variance of $x_i(t)$ associated with the kth component ρ_i^k.
- We can discard components with low variance, e.g. only keep the first K **principal components** such that $\sum_{k=1}^{K} \lambda_k \geq 0.95$, i.e. that represent 95% of the variance.

Note that the ρ_i^k is an orthogonal matrix that defines a new basis for the N-dimensional space such that the projections of $x_i(t)$ onto this new basis are uncorrelated with variance $\propto \lambda_k$.

Orthogonal means the matrix multiplied by its transpose gives the identity matrix $\mathbb{I}$:

$$\sum_i \rho_i^k \rho_i^l = \mathbb{I}_{kl} = \begin{cases} 0 \text{ if } k \neq l \\ 1 \text{ if } k = l \end{cases}$$

If we now project $x_i(t)$ onto the ρ_i^k, $x_i(t) = \sum_k a_k(t)\rho_i^k$, show that $a_k(t) = \sum_i \rho_i^k x_i(t)$ and now

$$\langle a_k(t)a_l(t)\rangle = \sum_{i,j} \rho_i^k \rho_j^l \langle x_i(t)x_j(t)\rangle = \sum_{i,j} \rho_i^k \Sigma_{ij} \rho_j^l = \sum_i \rho_i^k \mathrm{tr}(\Sigma)\lambda_l \rho_i^l = \mathrm{tr}(\Sigma)\lambda_k \mathbb{I}_{kl}$$

PCA as optimisation problem

We can also represent this procedure as an optimisation problem.

We define the projection of $x_i(t)$ onto some unit vector ρ_i^1 ($\rho^1 \cdot \rho^1 = 1$):

$$a_1(t) = x(t) \cdot \rho^1$$

We choose the ρ^1 such that the variance of $a_1(t)$:

$$\langle a_1(t)^2 \rangle = \langle (x(t) \cdot \rho^1)^2 \rangle$$

is maximised. This is an optimisation problem!

$$\max_{\{\rho_i^1\}} \sum_{i,j} \left\langle x_i(t) x_j(t) \rho_i^1 \rho_j^1 \right\rangle$$

such that

$$\sum_i \rho_i^1 \rho_i^1 = 1 \qquad \leftrightarrow \lambda_1$$

PCA as optimisation problem

KKT gives us from stationarity

$$0 = \frac{\partial \mathcal{L}}{\partial \rho_i^1} = \frac{\partial f}{\partial \rho_i^1} - \lambda_1 \frac{\partial g_1}{\partial \rho_i^1} = 2 \sum_j \rho_j^1 \langle x_i(t) x_j(t) \rangle - 2\lambda_1 \rho_i^1$$

This is nothing other than the eigenvalue equation for the covariance matrix $\Sigma_{ij} = \langle x_i(t) x_j(t) \rangle$!

Now consider the remainder defined by

$$\delta_i(t) = x_i(t) - a_1(t) \rho_i^1$$

Now let's find a second unit vector ρ_i^2 which is orthogonal to ρ_i^1 and points in the direction of greatest variance of the remainder $\delta_i(t)$

$$\max_{\{\rho_i^2\}} \sum_{i,j} \left\langle (\delta_i(t) \cdot \rho^2)^2 \right\rangle = \max_{\{\rho_i^2\}} \sum_{i,j} \left\langle (x_i(t) \cdot \rho^2)^2 \right\rangle$$

where we've used the fact that $\rho^1 \cdot \rho^2 = 0$. Repeating optimisation, we get another eigenvalue-eigenvector pair. Repeat until we have all eigenvalues and eigenvectors.

Application to Power System

Application to power injections for highly renewable European system

We're now going to apply PCA to the solved dispatch and network flows for a highly renewable European power system.

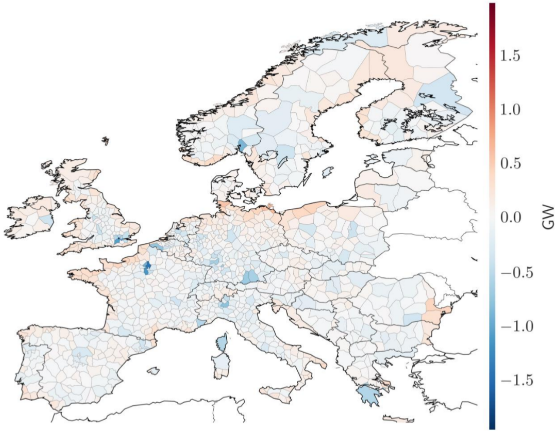
First we apply PCA to the power injections $p_i(t) = \sum_s g_{i,s}(t) - d_i(t)$ (generation minus demand). We compute the power injection covariance matrix:

$$\Sigma^p_{ij} = \langle p_i(t) p_j(t) \rangle - \langle p_i(t) \rangle \langle p_j(t) \rangle$$

NB: i, j run over the N different network nodes.

Next we find the eigenvectors and eigenvalues λ^p_k that represent the principal components.

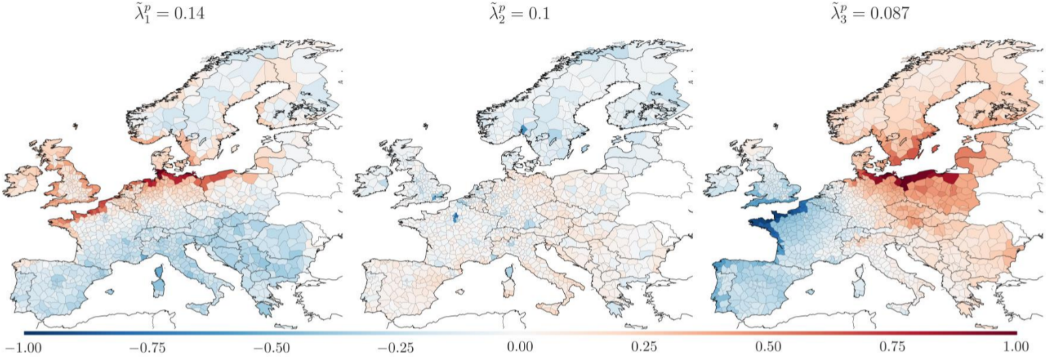
Average power injection $\langle p_i(t) \rangle$ at each node:



Source: Hofmann et al, 2018

Power injection components

The first 3 principal components represent the major axes of weather variations (1st is coastal wind production, 2nd is North-South seasonal pattern, 3rd is East-West load and solar daily pattern - check by Fourier transforming projection onto components):



Source: Hofmann et al, 2018

Application to resulting power flows

Next we apply PCA to the resulting power flows f_ℓ, related via the Power Transfer Distribution Factors

$$f_\ell = \sum_i H_{\ell i} p_i$$

(we use the notation $H = K^t L^{-1}$ for the PTDF to make things easier later).

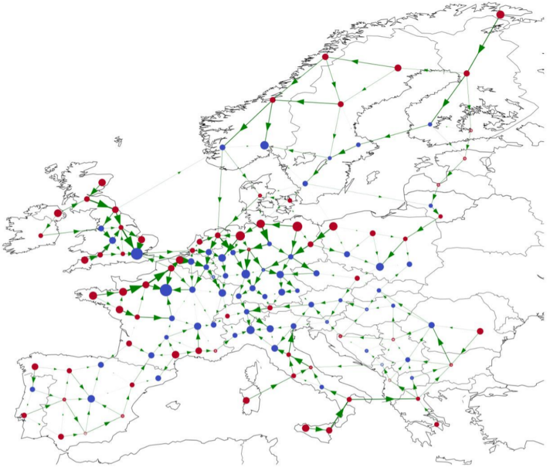
We compute the power flow covariance matrix:

$$\Sigma^f_{\ell m} = \langle f_\ell(t) f_m(t) \rangle - \langle f_\ell(t) \rangle \langle f_m(t) \rangle$$

NB: ℓ, m run over the L different network lines.

Next we find the eigenvectors and eigenvalues λ^f_n that represent the principal components.

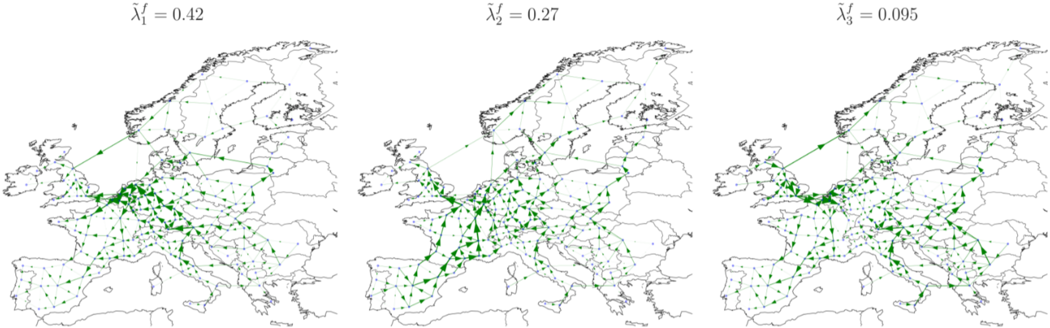
Average power flow $\langle f_\ell(t) \rangle$ at each line:



Source: Hofmann et al, 2018

Power flow components

The first 3 principal components represent the major flow (1st is flow to North-West, 2nd to North-East and 3rd shows multiple directions). Note that the first three make up a **much larger share of the variance** than for the power injection case.



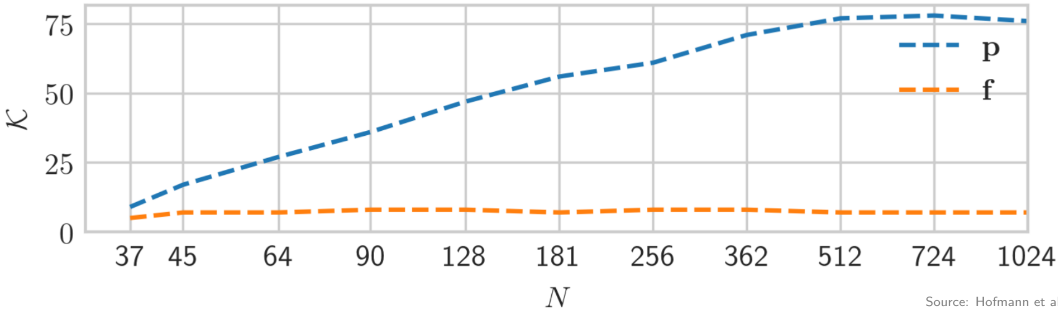
Source: Hofmann et al, 2018

Number of relevant components

- How many principal components K do we need to represent 95% of the total variance?

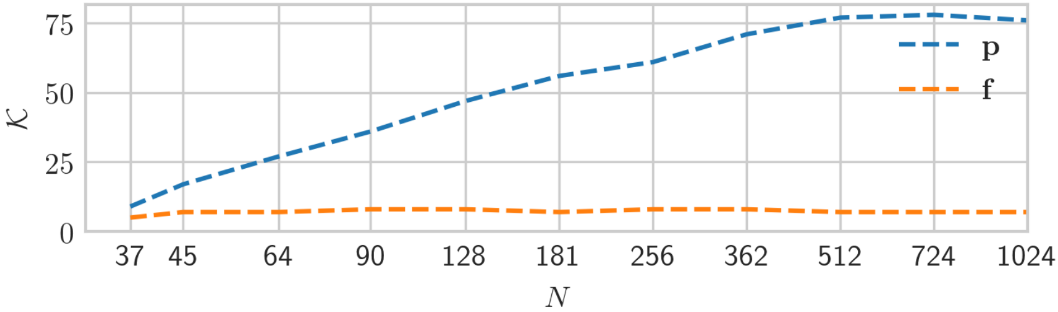
$$\sum_{k=1}^{K} \lambda_k \geq 0.95$$

- How does this number depend on the spatial resolution, i.e. the number of network nodes N used to represent the European grid?



Source: Hofmann et al, 2018

Number of relevant components



This graph is odd for (at least) 3 reasons:

- Why does the number of components required for the power injection rise then saturate at several hundred nodes? (Answer: correlation length)
- Why are so few components required to represent the power flow?
- Why doesn't the number of components change for the power flow?

Source: Hofmann et al, 2018

Relation of injection to flow covariance matrix

We have the following equations:

$$f_\ell = \sum_i H_{\ell i} p_i$$

$$\Sigma^p_{ij} = \langle p_i(t) p_j(t) \rangle - \langle p_i(t) \rangle \langle p_j(t) \rangle$$

$$\Sigma^f_{\ell m} = \langle f_\ell(t) f_m(t) \rangle - \langle f_\ell(t) \rangle \langle f_m(t) \rangle$$

So how are the flow covariance $\Sigma^f_{\ell m}$ and injection covariance Σ^p_{ij} matrices related?

Relation of injection to flow covariance matrix

We have the following equations:

$$f_\ell = \sum_i H_{\ell i} p_i$$

$$\Sigma^p_{ij} = \langle p_i(t) p_j(t) \rangle - \langle p_i(t) \rangle \langle p_j(t) \rangle$$

$$\Sigma^f_{\ell m} = \langle f_\ell(t) f_m(t) \rangle - \langle f_\ell(t) \rangle \langle f_m(t) \rangle$$

So how are the flow covariance $\Sigma^f_{\ell m}$ and injection covariance Σ^p_{ij} matrices related?

$$\Sigma^f = H \Sigma^p H^t$$

Relation of injection to flow covariance matrix

Now consider another $N \times N$ matrix M defined by

$$M = \Sigma^p H^t H$$

Note that the first term Σ^p comes from the injection pattern, whereas the second part $H^t H$ is entirely determined by the topology of the network (built from K and L).

If ν^k is an eigenvector of M with eigenvalue η_k, $M\nu^k = \eta_k \nu^k$, show that $H\nu^k$ is an eigenvector of Σ^f with eigenvalue η_k.

Relation of injection to flow covariance matrix

Now consider another $N \times N$ matrix M defined by

$$M = \Sigma^p H^t H$$

Note that the first term Σ^p comes from the injection pattern, whereas the second part $H^t H$ is entirely determined by the topology of the network (built from K and L).

If ν^k is an eigenvector of M with eigenvalue η_k, $M\nu^k = \eta_k \nu^k$, show that $H\nu^k$ is an eigenvector of Σ^f with eigenvalue η_k.

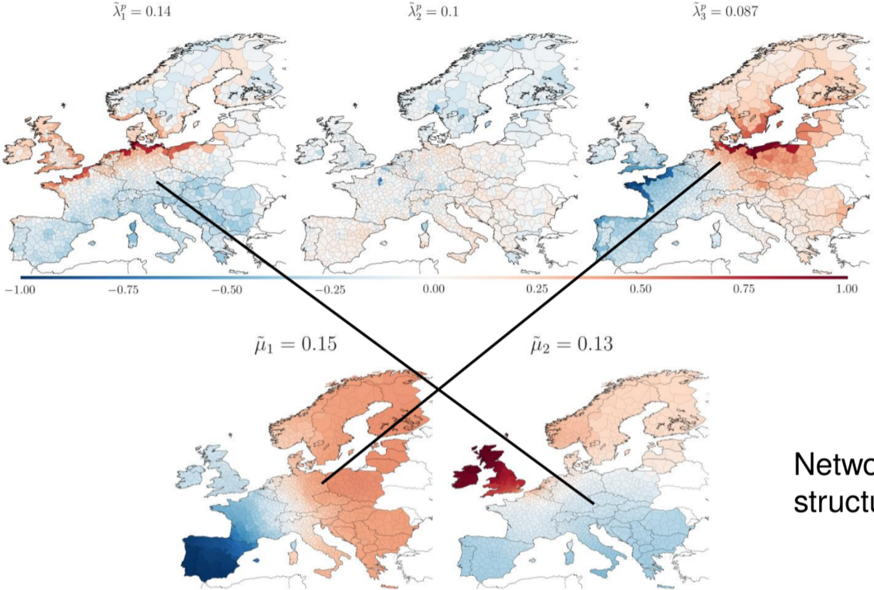
$$\Sigma^f H\nu^k = H\Sigma^p H^t H\nu^k = HM\nu^k = \eta_k H\nu^k$$

So to analyse the principal components of the flow, it suffices to study the eigenvectors of matrix M.

It turns out that if the first few eigenvectors of $H^t H$ and Σ^p with the strongest eigenvalues strongly overlap, then they magnify each other to the exclusion of other eigenvectors.

This is what happens for M (and by extension Σ^p): the eigenvectors of $H^t H$ magnify only the first few principal components of Σ^p, which then dominate Σ^f.

Network topology reinforces power injection pattern to magnify flow pattern



Injection pattern

Network structure

Network Topology reinforces flow pattern

To find out more, see our paper:

Fabian Hofmann, Mirko Schäfer, Tom Brown, Jonas Hörsch, Stefan Schramm, Martin Greiner, "Principal Flow Patterns across renewable electricity networks," EPL, 2018, **link**