Indices construction using linear and ordinal expert estimations

Vadim Strijov

Computing Center of the Russian Academy of Sciences

There are lots of ways to construct indices. However, when algorithms are chosen and some results obtained, the following question arises:

- **How to show adequacy of the calculated indices?**

To answer the question analysts invite experts. The experts express their opinion and then the second question arises:

- **How to show that expert estimations are valid?**

What is an index?

- There is a set of objects, i.e. power plants:
  - Beckjord
  - East Bend
  - Miami Fort
  - Zimmer
- Index a measure of an object's quality,
  a scalar corresponded to an object.
- Expert estimation of an object's quality
  could be an index, too.

How to construct an index?

- Assign a comparison criterion
- Gather a set of comparable objects
- Gather features of the objects
- Make a data table: objects/features, i.e.

#	Plant Name	Plant Type	Total Net Generation	CO_2 emission	NO_X emission	SO_X emission	Population density
			10^6 KWHours	Short tons per month	Short tons per month	Short tons per month	Qty per sq.mile
1	**Beckjord**	Coal	458505	191	16	45	23
3	**East Bend**	Coal	356124	147	16	43	34
4	**Miami Fort**	Coal	484590	204	6	23	45
5	**Zimmer**	Coal	818435	329	5	64	34
		Optimal value	max	min	min	min	min

The criterion could be: **Ecological footprint of a plant**

Notations

$A=\{a_{ij}\}$ – $(n\times m)$ real matrix, **data set**,

$\mathbf{q}=[q_1,\ldots,q_m]^{\mathrm{T}}$ – vector of **objects' indices**,

$\mathbf{w}=[w_1,..,w_n]^{\mathrm{T}}$ – vector of **features' importance weights**,

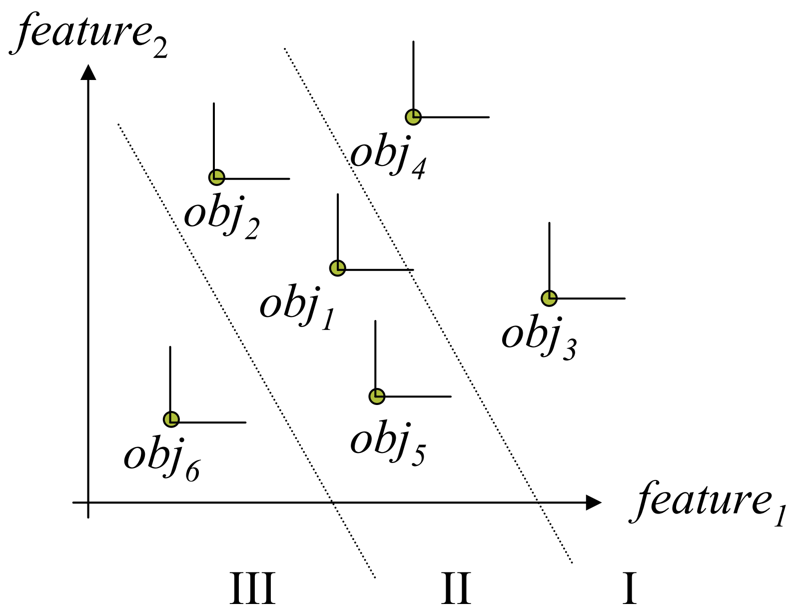
$\mathbf{q}_0$, $\mathbf{w}_0$ – expert estimations of indices and weights.

Usually, data prepared so that

1. the minimum of each feature equals 0, while the maximum equals 1;
2. the bigger value of each implies better quality of the index.

The first method, Pareto slicing

An easiest method to obtain indices in ordinal scales is to find non-dominated objects at each slicing level.

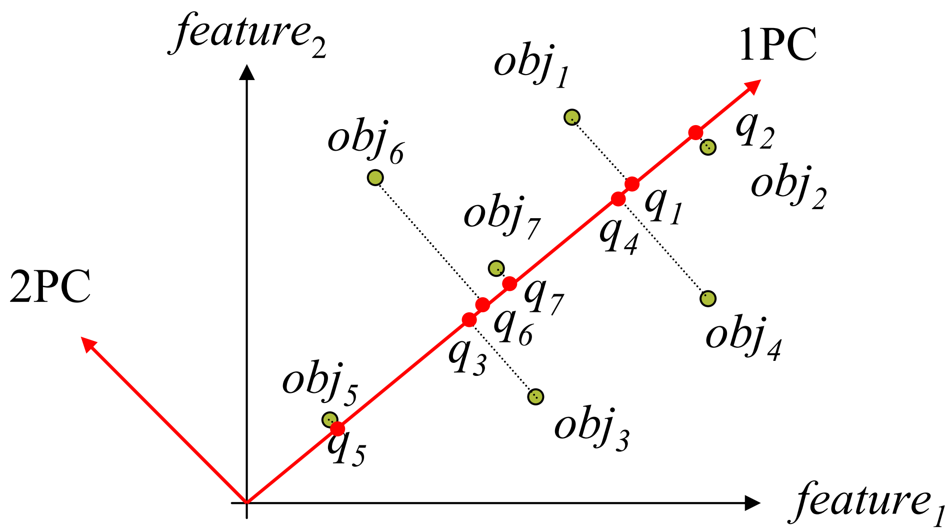


Unsupervised way,

Principal Components Analysis

Z=AW, where W—rotation matrix of the principal components.

$\mathbf{q}_{PCA}=A\mathbf{w}_{1PC}$, where $\mathbf{w}_{1PC}$ is the 1st column vector of W.



PCA gives minimal mean square error between objects and their projections.

Supervised way-1, the Weighted sum

$$\mathbf{q}_1 = A\,\mathbf{w}_0.$$

Supervised way-2,

the Expert-Statistical Technique

$\mathbf{w}_1 = \arg\min \|\mathbf{q}_0 - A\,\mathbf{w}_1\|$,

or

$\mathbf{w}_1 = A^+\,\mathbf{q}_0$.

$A^+ = VW^{-1}U^T$, while $A = UWV^T$ – singular values decomposition.

The problem of concordance

- **We have**

The data table A

Expert estimations $\mathbf{q}_0, \mathbf{w}_0$.

Calculated weights and indices $\mathbf{q}_1, \mathbf{w}_1$.

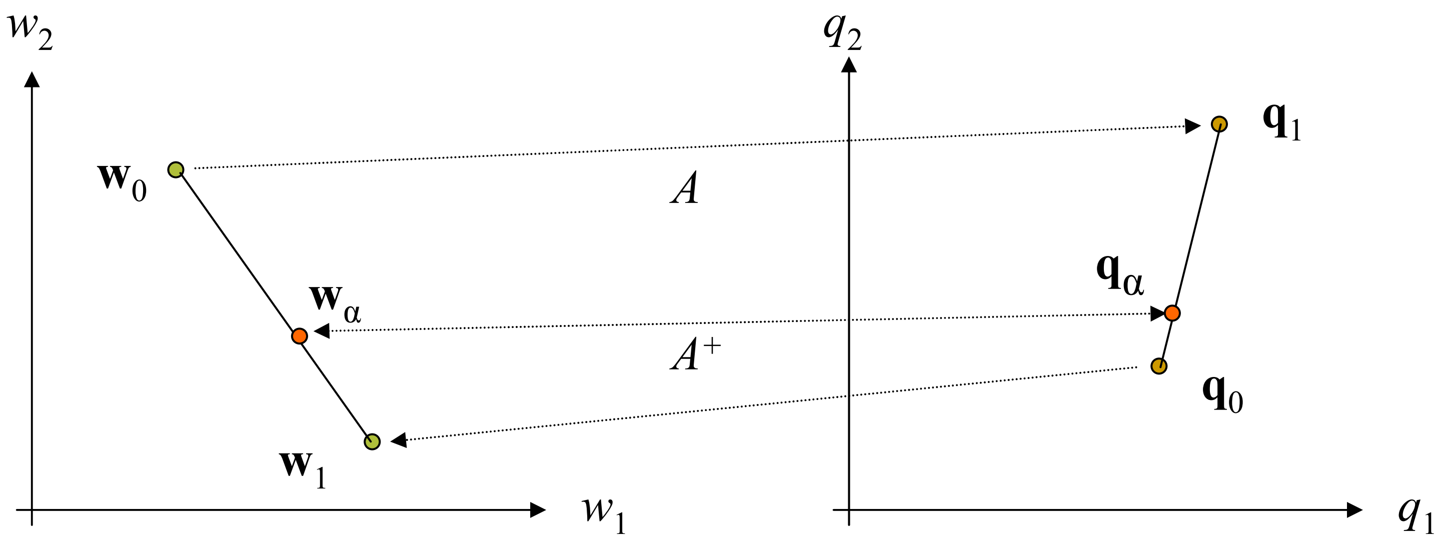
- **Dilemma**

Calculated indices are not the same as the expert estimations for the indices;

as well, calculated weights are not the same as the expert estimations of the weights:

in general, neither $\mathbf{q}_0 \neq A\mathbf{w}_0$, nor $\mathbf{w}_0 \neq A^+\mathbf{q}_0$.

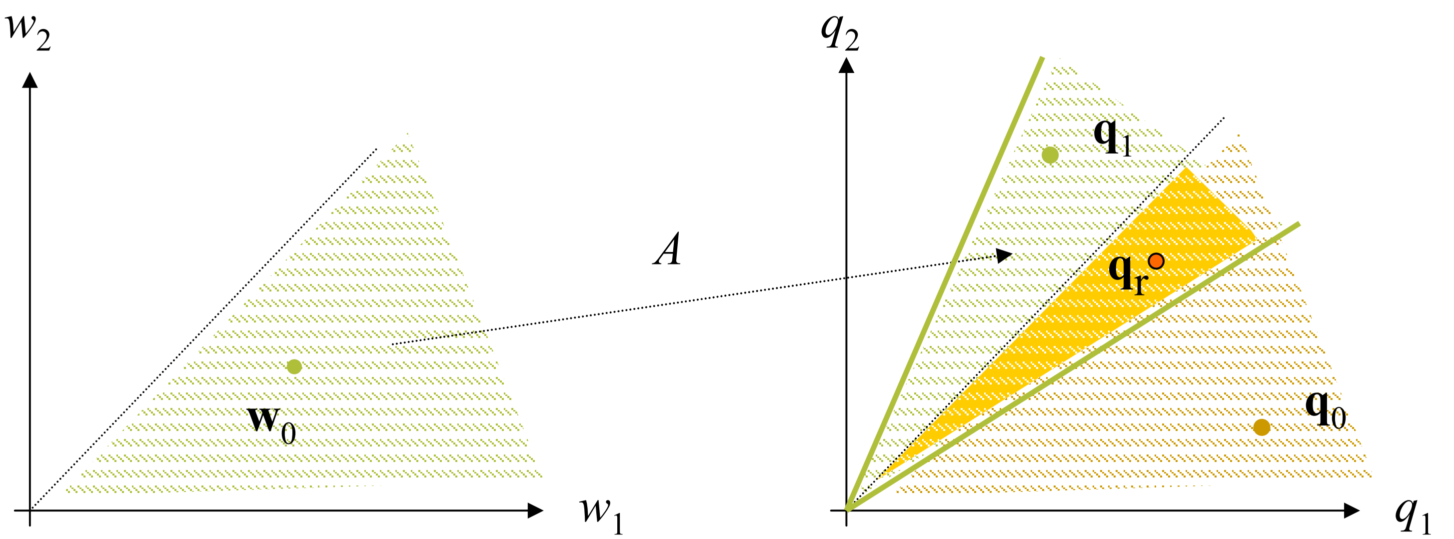
Linear concordance procedure



$$\mathbf{w}_\alpha = \alpha A^{+}\mathbf{q}_0 + (\mathbf{1}-\alpha)\mathbf{w}_0, \qquad \mathbf{q}_\alpha = (1-\alpha)A\mathbf{w}_0 + \alpha\mathbf{q}_0.$$

Parameter α is in [0,1].
α = 0, then we trust expert estimations of the weights
α = 1, then we trust expert estimations of the indices

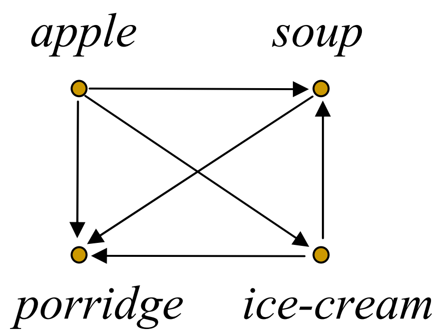
Ordinal concordance procedure

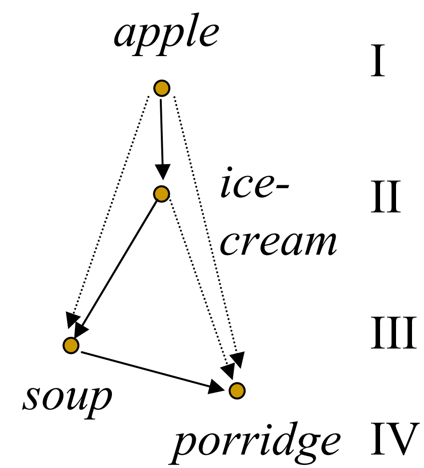


$$\mathbf{w}_0 = [w_1 > w_2 > \ldots > w_n]^{\mathrm{T}}, \quad \mathbf{q}_0 = [q_1 > q_2 > \ldots > q_m]^{\mathrm{T}}.$$

Pair-wise comparison

	a	*s*	*p*	*i-c*
apple	●	+	+	+
soup		●	+	–
porridge			●	–
ice-cream				●





If an object in a row is better than the other one in a column then put “+”, otherwise “-”.

Make a graph, *row* + *column* means *row* ●→● *column*.

Find the top and remove extra nodes.

Results of the concordance are

- adequate indices.
- reasoned expert estimations.
- We know why expert valued an object
- and what contribution a feature makes to an index.

And we have precise weights

to make indices in the future.

Thanks!

Vadim Strijov, strijov@ccas.ru