Chapter 16
Partial Order and Related Disciplines

16.1 Partial Order and Mathematics

It is always hard to try a positioning. Nevertheless, it may help interested readers to find their way through the jungle of concepts, relations, and equations of this text.

Certainly, partial order has to do with graph theory in discrete mathematics, as its visualization is a digraph and questions like connectivity or identification of articulation points and of separated subsets are typical of graph theory; see, e.g., Wagner and Bodendiek (1989) and Patil and Taillie (2004). There is also a connection to the network domain, as partial order constitutes a directed graph, which is one of the characteristics of networks. In our applications here, there is always a matrix, which quantifies the multi-indicator system, the data matrix. With or without the interim step of deriving the rank matrix, we arrive at a partial order. Once, however, the poset is derived, it may be analyzed as a mathematical object on its own right. So the comparative structure is not only inherent in the rank matrix, but other matrices can be found, which describe the order relation in different ways. Such matrices together with their arithmetic are realizations of incidence algebra; see Stanley (1986). So we have one example of relations to algebra. Another relation to algebra is the algebra of posets, i.e., how to combine them or how to find simpler graphs whose combination by different well-defined operations leads back to the original digraph (Neggers and Kim, 1998). The dismantling of an empirical poset finds its limits as to how far we can find an interpretation of the resulting simpler digraphs. Recently, an important step was done by partitioning the attribute set, construction of the posets induced by the indicators' subsets, and by an appropriate recombination of the posets. This procedure allows us to consider preferences among the attributes (Rademaker et al., 2008). Another relation to algebra can be established, when we realize that formal concept analysis is a powerful tool based on partial order. Concepts, being ordered by inclusion and forming a lattice, can be related to each other by two algebraic operations. Finally, formal concept analysis can be analyzed in terms of universal algebra.

The study of linear extensions and their properties, see, e.g., Habib et al. (1988) and Syslo (1985), is a field of combinatorics in discrete mathematics (see also Chapter 18), which also is important in applications of partial order analyzed in this monograph.

R. Brüggemann, G.P. Patil, *Ranking and Prioritization for Multi-indicator Systems*, Environmental and Ecological Statistics 5, DOI 10.1007/978-1-4419-8477-7_16,

16.2 Partial Order and Statistics

Welzl et al. (2002) have characterized the analysis of a data matrix with concepts of variance, distance, and order. This monograph is dedicated to partial order analysis (PoA) concerned with extracting ordinal information from the data matrix, when the aim of the ranking within a multi-indicator system is known and the indicators are appropriately oriented (Chapter 3). With statistical data processing of the day, we are confronted with three major questions:

Table 16.1 Comparison of partial order analysis (PoA), principal component analysis (PCA), and cluster analysis (CA)

Method	Advantage	Disadvantage	Remark
Partial order analysis (PoA)	Comparison of objects is the central topic. One of the most important tools of PoA is to provide a linear order of the objects without the need of a weighting of the indicators	We miss significance tests	An important concept is the ordinal modeling (Chapter 6). Its aim is to extract that information from the data matrix which is relevant for comparisons. However, we still miss theoretical guidance
Principal component analysis (PCA)	Insight into the structure of data matrix. Visualization techniques like biplots. Test statistics available. Weights can be constructed with which superindicators (“pillars”) can be constructed	No direct access to order. Assumes linear models	Techniques are available to extend ordination technique to nonlinear models
Cluster analysis (CA)	Powerful access to group items due to their similarity. The items may be objects or attributes of the data matrix	Methodological artifacts, like the chain effect in single linkage methods	Reordering can be introduced and visualized as within any partitioning, there is freedom to order the classes of objects according to the partial order relations among them; see Mucha (2002)

1. How do we assess significance of partial order results when there is uncertainty in the data?
2. What kind of relation between partial order methods and multivariate statistics tools can be established?
3. How can we find good exploration and visualization techniques beyond Hasse diagram?

Ad 1: Assessing the significance of partial order results is done by simulations and can be a good topic for future research in partial order; see in that context Sørensen et al. (1998, 2000, 2009), Saltelli et al. (2008), Saltelli and Annoni (2010), and Annoni et al. (2012).

Ad 2: The second question may be answered by checking what the standard methods of multivariate statistics offer (Table 16.1).

Ad 3: Partially, this question is answered by application of statistical techniques (Table 16.1); however, the background of this question is "data mining" which is more closely discussed in Section 16.4.

16.3 Partial Order and Fuzzy Concepts

In Chapter 6, we applied the concept "ordinal modeling" because the order relations should not be overburdened by small data differences, which nevertheless are ordinal interpreted. One possibility is to introduce fuzzy concepts into partial order analysis. There are several possibilities:

(a) Instead of a hierarchical clustering, we may use a fuzzy cluster concept. This at least avoids the hard decision of whether or not an object belongs to a cluster. After an appropriate selection of a cluster center, partial order is applied to the cluster centers as fictitious objects; see Luther et al. (2000).

(b) In the Kosko approach, the crisscrossing of data profiles is replaced by a fuzzy subsethood. We have described this procedure and shown applications in this monograph (Chapters 6 and 11). Is the Kosko measure the ideal measure to obtain a fuzzy subsethood? What can be said about its relation to the mutual probabilities, derived from linear extensions? Could we find better measures than the Kosko measure which sums the indicator values? How to find optimal α-cuts (optimal with respect to what)? The question of α-cuts is discussed by Annoni et al. (2008). There, however, is still a need for theoretical work.

(c) An interesting alternative is provided by Fattore (2008), who does not use the Kosko measure but the height of an object in the set of linear extensions as a fuzzy membership function. This attractive method needs further attention, especially in its theoretical implications. It does not, however, directly contribute to "ordinal modeling."

16.4 Partial Order and Data Mining

16.4.1 Overview

We come back to some of the aforementioned points: There is a limit at present in the graphical presentation of partial order. Hasse diagrams are a wonderful visualization tool for working with posets with few elements. Alternatively, one may look for other visualization techniques, POSAC (Chapter 3), Bertin strategy, RRR, as well as the elimination of endmembers (EoE), and the use of dominance diagrams (Chapter 5).

16.4.2 POSAC

The idea is to project the m-dimensional attribute space due to m attributes of the information base into a two-dimensional plane spanned by two latent order variables LOV(1) and LOV(2). The projection is done by keeping at maximum the comparabilities found in the m-dimensional property space. The scatter plot allows an inspection of the whole object set and is – in principle – not restricted by the number of objects. The interpretation of the latent variables is sometimes difficult, and the computational methods lead to approximate solutions. By partial order dimension analysis, it can be checked as to whether a two-dimensional representation is intrinsically possible. However, there are still computational problems to be mastered.

16.4.3 Bertin Strategy

The starting point of Bertin strategy is the multicoordinate approach: The ranks of objects and the labels of the attributes represent the rows and columns, resp., of the Bertin matrix. An iterative procedure of row and column permutations is applied until the matrix gets as far as possible a blocked form (see http://en.wikipedia.org/wiki/Jacques_Bertin). The homogeneity of the block matrices is checked in order to maximize the homogeneity within the blocks and to minimize it among the block matrices. The complexity of this procedure corresponds to that of traveling sales man problem, hence heuristic approaches (cluster analysis) are to be applied. According to Welzl et al. (2002) there is, however, still need of statistical tests.

16.4.4 Rank Range Run (RRR) and Elimination of Endmembers (EoE)

The RRR approach was introduced by Myers et al. (2006). The main steps in rank range run (RRR) are the following:

1. Assign ranks to each of the attributes.
2. Order the objects according to their minimum rank.
3. If there are equal minimum ranks, take the maximum rank as an ordering property; if the maximal rank is equal, then take the median.
4. Each object is represented by a vertical line according to its minimum and maximum rank values.
5. Locate the objects along the horizontal axis by following the order of Step 3.
6. Mark the median of the ranks for each object.

Depending on how the orientation is selected (rank = 1 for the best or the worst value) the graphical display gives an overview about the ordinal properties with a focus on the whole set and with less focus on a single object. The length of any line associated with an object is associated with a degree of inequality. Here we stress that this approach does not depend on the representation by a graph, like the Hasse diagrams, and is suitable for data mining. Elimination of endmembers in RRR can be seen as a purification process; see Myers et al. (2006): One may ask oneself as to which elimination of an object with the highest rank gives the most reduction in the rank range (EoE gets positive values) and which object with minimum rank would give the most reduction in the rank range (EoE gets negative values). One may plot the EoE values vs the rank range run. This kind of analysis helps to identify attributes whose improvement would give a high effect on the overall status of an object or attributes whose status is responsible for the bad evaluation of some objects. For an example, see Newlin and Patil (2010).

Graphical analysis tools based on partial order concepts can be further developed, mainly based on the dominance relation (Chapter 5) and the local partial order approach (Chapter 9); see Myers and Patil (2010).

16.5 Partial Order and Network Analysis

Network is a topic in the systems analysis research because the analysis of networks as an abstract concept can deliver important results in several fields, such as

- Biology: food webs, biochemical networks
- Sociology: social networks
- Chemistry: reaction networks inclusive of biochemical networks
- Engineering: electrical networks
- Information technology: dependency networks of, e.g., compilers
- Urban drinking water systems
- River networks
- Transportation networks
- Bridge networks
- Implication networks as provided by the formal concept analysis (see Chapters 8, 13 and 15).

Even partial order can be considered as networks as comparative evaluation networks.

(a) What can network analysis learn from partial order approach and (b) what tools of network analysis can be helpful for partial order theory?

(a) Networks are vertex/node and directed edge evaluated graphs. So they fall into the general domain of graph theory. An obvious introduction of partial order is possible by considering the set of all subgraphs of a given graph and to order them by the (graph theoretical) inclusion relation. In chemistry, this conceptualization of molecular graphs and subgraphs as posets made an important contribution to the discovery of the relationships of chemical properties with molecular graphs obtained from molecules. See Klein (1986) in this context and Klein and Bytautas (2000) and Ivanciuc et al. (2005, 2006a, b) for the extended concepts like reaction graphs and substitution graphs.

Another way to analyze a number of networks by partial order methods comes when networks are to be compared, i.e., when the set of networks is a poset. Then according to the aim of the comparative study, the indicators are obtained from the characterization of networks by network invariants, such as diameter, number of nodes, eccentricity, and centrality. In this case, networks are considered as objects described by multiple indicators. We have not seen this promising approach in the network analysis literature so far.

Partial order finds application in network analysis when the set of vertices can be partitioned into two classes. In that case a Galois lattice provides useful tools (Wasserman and Faust, 2009).

(b) A powerful use of network concepts comes into play with well-known software graphviz to which PyHasse has an interface. The software graphviz resulted from network analysis and its graphical representation. The software graphviz (Gansner et al., 2009) uses many theoretical results concerning directed graphs and allows the user to draw Hasse diagrams under different options, such as

- Rank separation,
- Merging parallel connecting lines,
- Aggregation of objects to clusters, and
- Constrain rank assignments (subgraphs may get their own location for sources and sinks).

References

Annoni, P., Fattore, M. and Bruggemann, R. (2008). Analyzing the structure of poverty by fuzzy partial order. In J. Owsinski and R. Bruggemann (Eds.), Multicriteria ordering and ranking: Partial orders, ambiguities and applied issues (pp. 107–124). Warsaw: Systems Research Institute Polish Academy of Sciences.

Annoni, P., Fattore, M. and Bruggemann, R. (2012). A multi-criteria fuzzy approach for analyzing poverty structure. Statistica & Applicazioni, accepted.

Bertin strategy. http://en.wikipedia.org/wiki/Jacques_Bertin.

Fattore, M. (2008). Hasse diagrams, poset theory and fuzzy poverty measures. Riv. Int. Sci. Soc., 1, 63–75.

Gansner, E.R., Koutsofios, E. and North, S. (2009). Drawing graphs with *dot*. http://www.graphviz.org/pdf/dotguide.pdf.

Habib, M., Moehring, R.H. and Steiner, G. (1988). Computing the bump number is easy. Order, 5, 107–129.

Ivanciuc, T., Ivanciuc, O. and Klein, D.J. (2005). Posetic quantitative superstructure/activity relationships (QSSARs) for chlorobenzenes. J. Chem. Inf. Model., 45, 870–879.

Ivanciuc, T., Ivanciuc, O. and Klein, D.J. (2006a). Modeling the bioconcentration factors and bioaccumulation factors of polychlorinated biphenyls with posetic quantitative super-structure/activity relationships (QSSAR). Mol. Divers., 10, 133–145.

Ivanciuc, T., Ivanciuc, O. and Klein, D.J. (2006b). Prediction of environmental properties for chlorophenols with posetic quantitative super-structure/property relationships (QSSPR). Int. J. Mol. Sci., 7, 358–374.

Klein, D.J. (1986). Chemical graph-theoretic cluster expansions. Int. J. Quant. Chem., 20, 153–171.

Klein, D.J. and Bytautas, L. (2000). Directed reaction graphs as posets. Match, 42, 261–290.

Luther, B., Bruggemann, R. and Pudenz, S. (2000). An approach to combine cluster analysis with order theoretical tools in problems of environmental pollution. Match, 42, 119–143.

Mucha, H. (2002). Clustering techniques accompanied by matrix reordering techniques. In K. Voigt and G. Welzl (Eds.), Order theoretical tools in environmental sciences – order theory (Hasse diagram technique) meets multivariate statistics (pp. 129–140). Aachen: Shaker-Verlag.

Myers, W.L. and Patil, G.P. (2010). Partial order and rank range runs for compositional complexes in landscape ecology and image analysis, with applications to restoration, remediation, and enhancement. Environ. Ecol. Stat., 17, 411–436.

Myers, W.L., Patil, G.P. and Cai, Y. (2006). Exploring patterns of habitat diversity across landscapes using partial ordering. In R. Bruggemann and L. Carlsen (Eds.), Partial order in environmental sciences and chemistry (pp. 309–325). Berlin: Springer.

Neggers, J. and Kim, H.S. (1998). Basic posets. Singapore: World Scientific Publishing.

Newlin, J.T. and Patil, G.P. (2010). Application of partial order to bridge engineering, stream channel assessment, and infra structure management. Environ. Ecol. Stat., 17, 437–454.

Patil, G.P. and Taillie, C. (2004). Multiple indicators, partially ordered sets, and linear extensions: Multi-criterion ranking and prioritization. Environ. Ecol. Stat., 11, 199–228.

Rademaker, M., De Baets, B. and De Meyer, H. (2008). New operations for informative combinations of two partial orders with illustrations on pollution data. Comb. Chem. High Throughput Screen., 11(9), 745–755.

Saltelli, A. and Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. Environ. Model. Softw., 25, 1508–1517.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., et al. (2008). Global sensitivity analysis – The primer. Chichester: Wiley.

Sørensen, P.B., Bruggemann, R., Thomsen, M., Gyldenkaerne, S. and Kjaer, C. (2009). How to guide and assess risk reduction using risk characterization indicators. Am. J. Appl. Sci., 6(6), 1255–1263.

Sørensen, P.B., Mogensen, B.B., Carlsen, L. and Thomsen, M. (2000). The influence on partial order ranking from input parameter uncertainty – Definition of a robustness parameter. Chemosphere, 41, 595–600.

Sørensen, P.B., Mogensen, B.B., Gyldenkaerne, S. and Rasmussen, A.G. (1998). Pesticide leaching assessment method for ranking both single substances and scenarios of multiple substance use. Chemosphere, 36(10), 2251–2276.

Stanley, R.P. (1986). Enumerative combinatorics (Vol. I). Monterey: Wadsworth and Brooks/Cole.

Syslo, M.M. (1985). A graph-theoretic approach to the jump-number problem. In I. Rival (Ed.), Graphs and order (pp. 185–215). Dordrecht: D. Reidel Publishing Company.

Wagner, K. and Bodendiek, R. (1989). Graphentheorie I: Anwendungen auf Toplogie, Gruppentheorie und Verbandstheorie. Mannheim: BI Wissenschaftsverlag.

Wasserman, S. and Faust, K. (2009). Social network analysis – Methods and applications. Cambridge: Cambridge University Press.

Welzl, G., Voigt, K. and Bruggemann, R. (2002). Order theory meets statistics – Environmental statistics. In K. Voigt and G. Welzl (Eds.), Order theoretical tools in environmental sciences – order theory (Hasse diagram technique) meets multivariate statistics (pp. 41–54). Aachen: Shaker-Verlag.