

The background of the slide is a complex, abstract composition. It features a central white rectangular area containing the title. Surrounding this area are various geometric and data-related patterns. On the left, there's a vertical strip with a grid of small '+' symbols. To its right, a network of red lines connects numerous green dots, resembling a graph or a spatial partitioning structure. Further right, there's a pattern of thin, light-colored lines forming a mesh. The overall color palette is muted, with shades of brown, beige, and light blue, giving it a technical and academic feel.

Basic Concepts of Partitioning Algorithms

Partitioning Algorithms: Basic Concepts

- Partitioning method: Discovering the groupings in the data by optimizing a specific objective function and iteratively improving the quality of partitions
- *K*-partitioning method: Partitioning a dataset D of n objects into a set of K clusters so that an objective function is optimized (e.g., the sum of squared distances is minimized, where c_k is the centroid or medoid of cluster C_k) c_k 是类 C_k 的中心点.
- A typical objective function: **Sum of Squared Errors (SSE)**
$$SSE(C) = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - c_k\|^2$$
- Problem definition: Given K , find a partition of K clusters that optimizes the chosen partitioning criterion
 - Global optimal: Needs to exhaustively enumerate all partitions
 - Heuristic methods (i.e., greedy algorithms): *K-Means, K-Medians, K-Medoids*, etc.
Local maximum.