specific order among the groups is reported (Non-Patent Document 1).

[0007] Embodiments of the present invention provide a data analysis apparatus, method, and program which can perform various data analysis by taking into account a group order among statistical samples.

MEANS FOR SOLVING THE PROBLEMS

[0008] According to a first aspect of the present invention, there is provided a data analysis apparatus which performs multivariate analysis with a plurality of data items on a plurality of statistical samples, the data analysis apparatus including a storage which records statistical data for managing the plurality of data items per statistical sample, and group information indicating an arrangement order of groups formed by the plurality of statistical samples; and a controller which performs predetermined calculation processing based on the statistical data and the group information, wherein the controller calculates, based on the statistical data, a kernel matrix in which a matrix element represents a predetermined relationship between a statistical sample corresponding to a row number of the matrix element in the plurality of statistical samples and a statistical sample corresponding to a column number of the matrix element, and performs calculation processing based on a partial least squares method under a predetermined condition defined by the kernel matrix and the group information, to calculate a score for the statistical samples.

[0008a] According to a second aspect of the present invention, there is provided a data analysis method for causing a computer to perform multivariate analysis with a plurality of data items on a plurality of statistical samples, wherein a storage of the computer records statistical data for managing the plurality of data items per statistical sample, and group information indicating an arrangement order of groups formed by the plurality of statistical samples, and the data analysis method includes, by the computer: calculating, based on the statistical data, a kernel matrix in which a matrix element represents a predetermined relationship between a statistical sample corresponding to a row number of the matrix element in the

plurality of statistical samples and a statistical sample corresponding to a column number of the matrix element; and performing calculation processing based on a partial least squares method under a predetermined condition defined by the kernel matrix and the group information, to calculate a score for the statistical samples.

[0008b] According to a third aspect of the present invention, there is provided a program for causing a computer to execute the above data analysis method.

EFFECTS OF THE INVENTION

[0009] The data analysis apparatus according to embodiments of the present invention can perform integrated analysis and nonlinear analysis on various types of statistical data based on the kernel matrix while reflecting a group order in scores based on group information. Consequently, it is possible to perform various types of data analysis while taking into account the group order among statistical samples.

BRIEF DESCRIPTION OF DRAWINGS

[0010]

Fig. 1A and 1B are views for explaining an outline of a data analysis method according to a first embodiment.

Fig. 2 is a chart illustrating metabolome data of a liver sample.

Fig. 3 is a chart illustrating metabolome data of a heart sample.

Fig. 4 is a chart illustrating metabolome data of a brain sample.

Fig. 5 is a chart illustrating metabolome data of a plasma sample.

Fig. 6 is a block diagram illustrating a configuration of a data analysis apparatus according to the first embodiment.

Fig. 7 is a flowchart illustrating data analysis processing by the data analysis apparatus.

Figs. 8A to 8F are expressions for explaining the