CdL in Scienze Statistiche ed Economiche - Università degli Studi di Milano-Bicocca

Esercitazione: La matrice dei dati

Esercitatrice: Chiara Gaia Magnani

Example 0.1. E' stato accennato a lezione che utilizzare il vettore delle mediane (ovvero ciascuna componente del vettore rappresenta la mediana della rispettiva variabile) non é una buona misura per rappresentare il baricentro di dati multidimensionali.

Infatti potrebbe succedere che questo vettore si collochi sul bordo dell'involucro convesso (convex hull) comprendente tutte le osservazioni.

Intuitivamente, l'involucro convesso di un insieme di punti è la forma che assumerebbe un elastico allargato in modo da contenere tutti i punti e poi lasciato libero di restringersi: un poligono che ha alcuni di quei punti come vertici e li contiene tutti.

Si costruisca un esempio di matrice di dati X con due variabili (p=2) e almeno cinque osservazioni $(n \ge 5)$ tale per cui il vettore delle mediane giace sul bordo dell'involucro convesso.