

## Esercitazione : La matrice dei dati

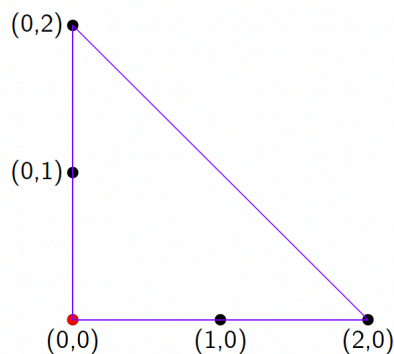
Esercitatrice: Chiara Gaia Magnani

**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.*

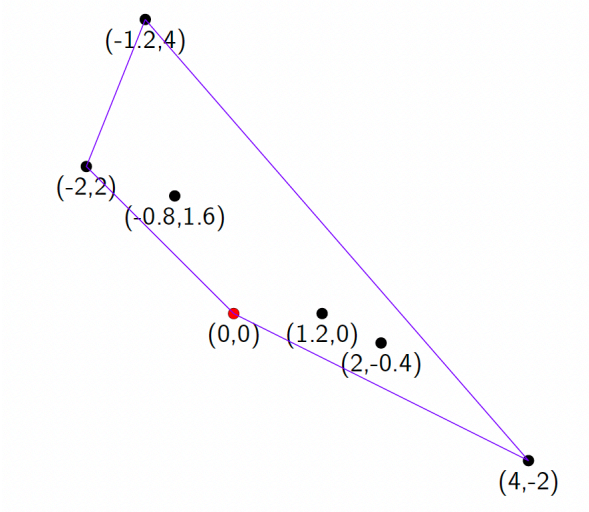
*Si costruisca un esempio di matrice di dati  $X$  con due variabili ( $p = 2$ ) e almeno cinque osservazioni ( $n \geq 5$ ) tale per cui il vettore delle mediane giace sul bordo dell'involucro convesso.*



*Dimostrazione.*

La mediana della prima variabile è  $(0, 0, 0, 1, 2) = 0$  e la mediana della seconda variabile è  $(0, 1, 2, 0, 0) = 0$ , quindi il vettore delle mediane è  $(0, 0)$  (rappresentato dal punto rosso) e giace sul bordo dell'involucro convesso che comprende i dati (rappresentato in blu).

Un altro esempio è il seguente:



□