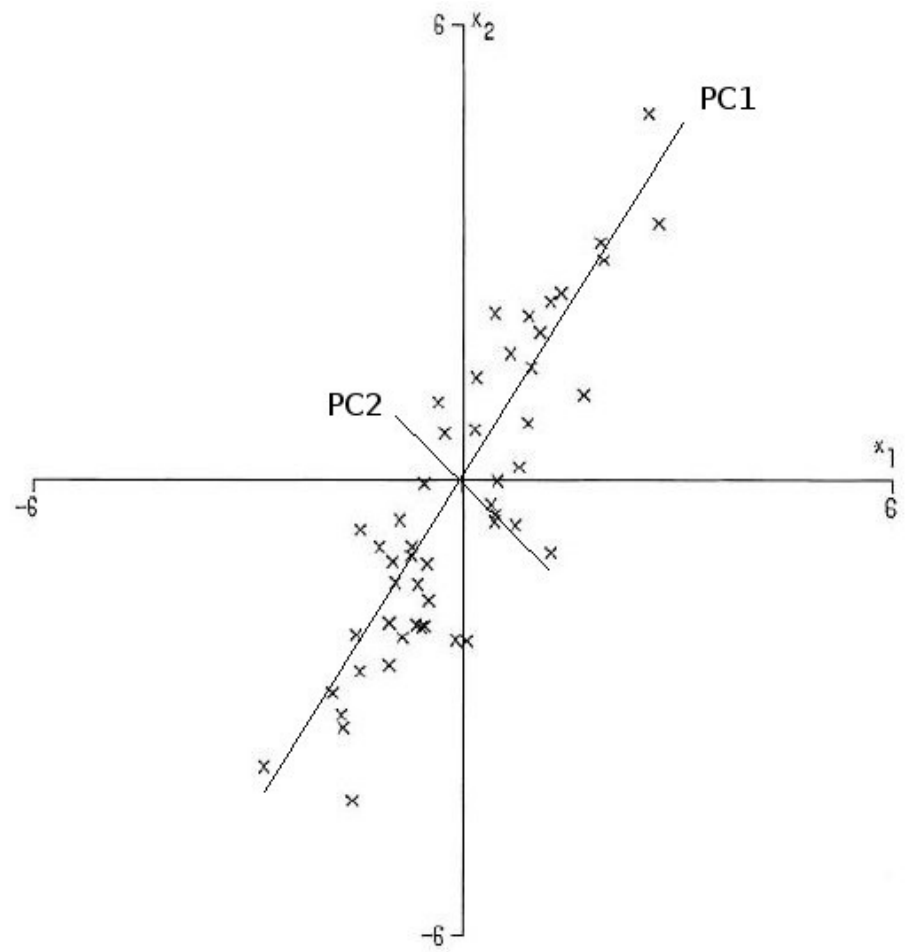
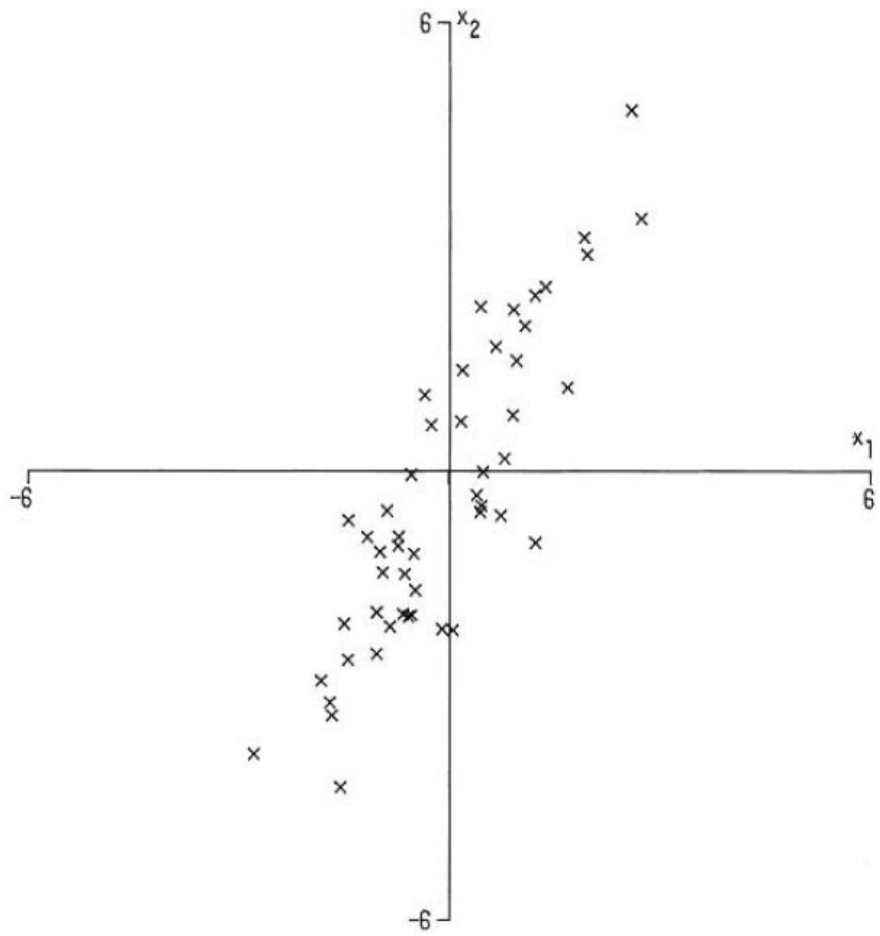
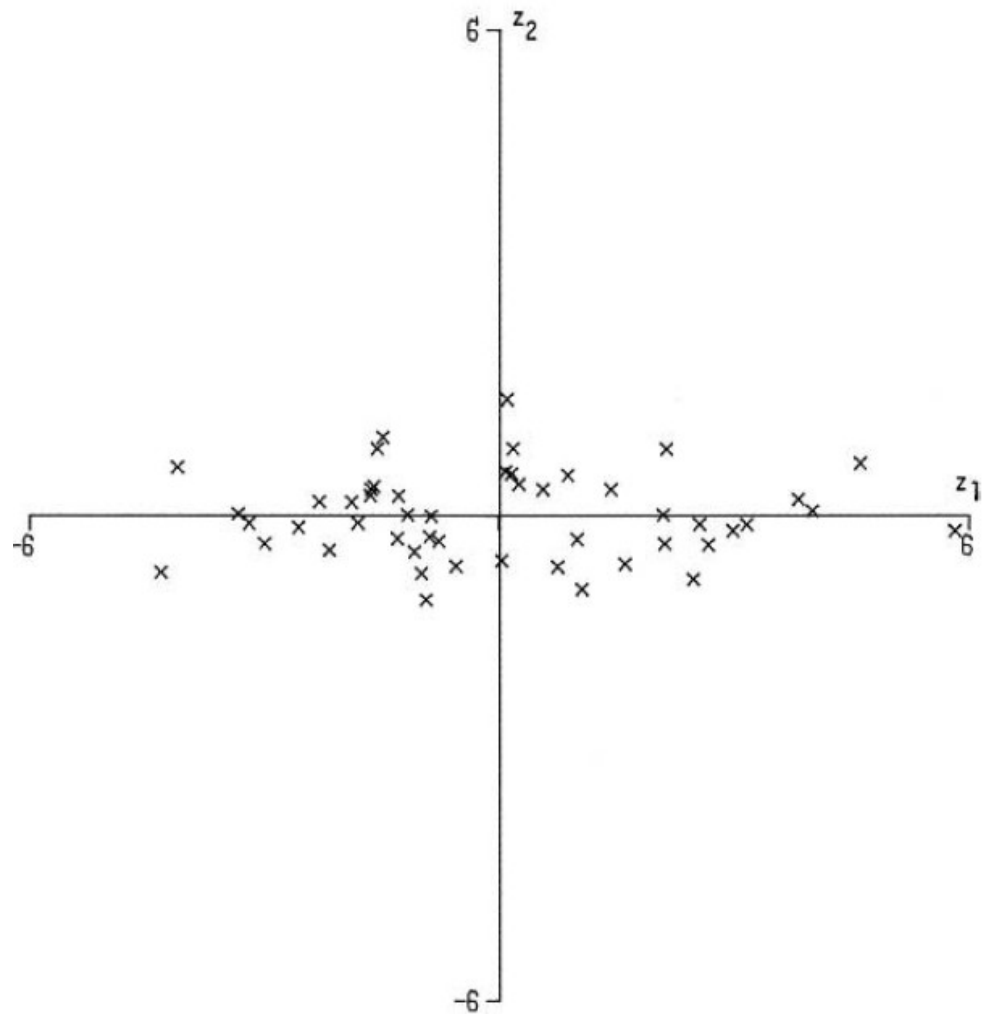


Ανάλυση Κύριων Συνιστωσών (PCA)

Σκοπός της μεθόδου η μείωση της διάστασης ενός συνόλου δεδομένων το οποίο αποτελείται από μεγάλο πλήθος αλληλοεξαρτώμενων μεταβλητών.



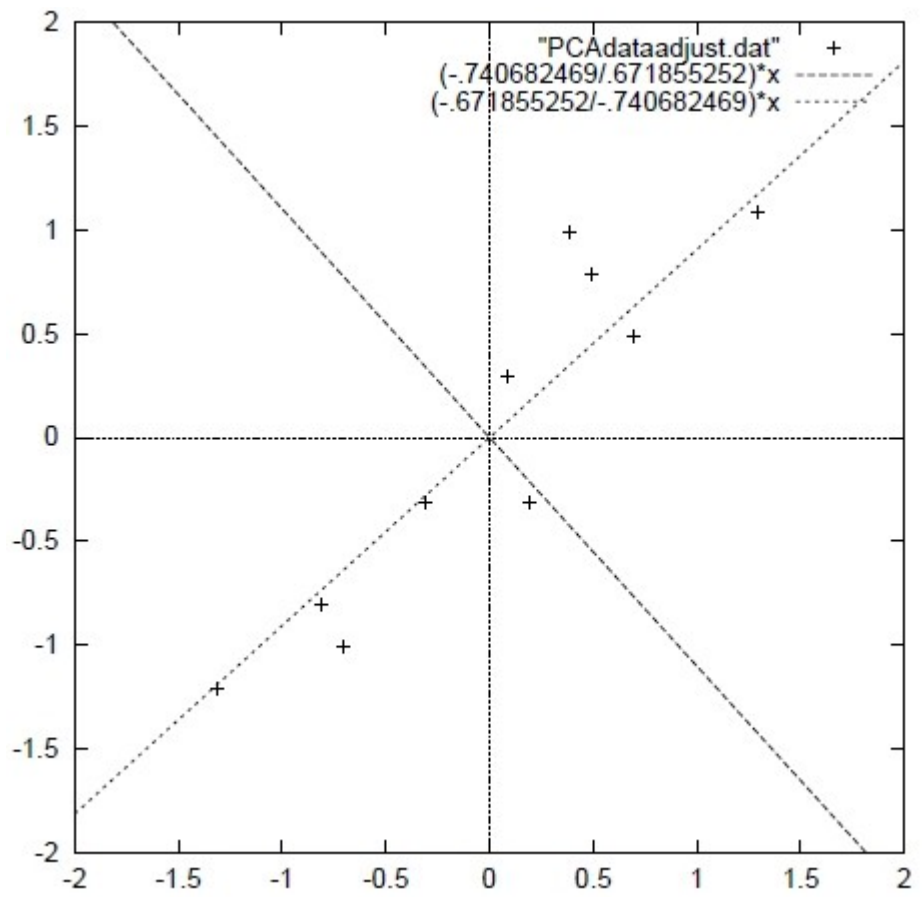
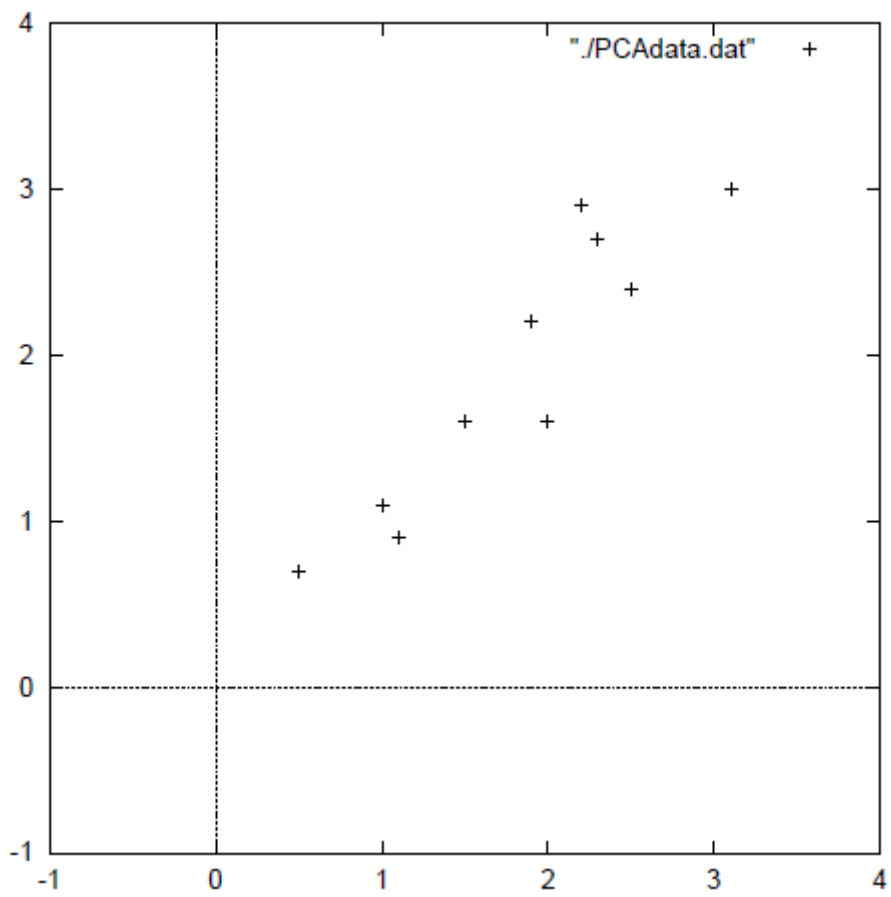


Για τον υπολογισμό των κύριων συνιστωσών του συνόλου δεδομένων:

- Υπολογισμός μητρώου συνδιασπορών
- Υπολογισμός των ιδιοτιμών του μητρώου, ταξινόμησή τους κατά φθίνουσα σειρά
- Εύρεση των αντίστοιχων ιδιοδιανυσμάτων που αποτελούν τον καινούριο χώρο
- Προβολή του συνόλου δεδομένων στον ιδιόχωρο

Παράδειγμα:

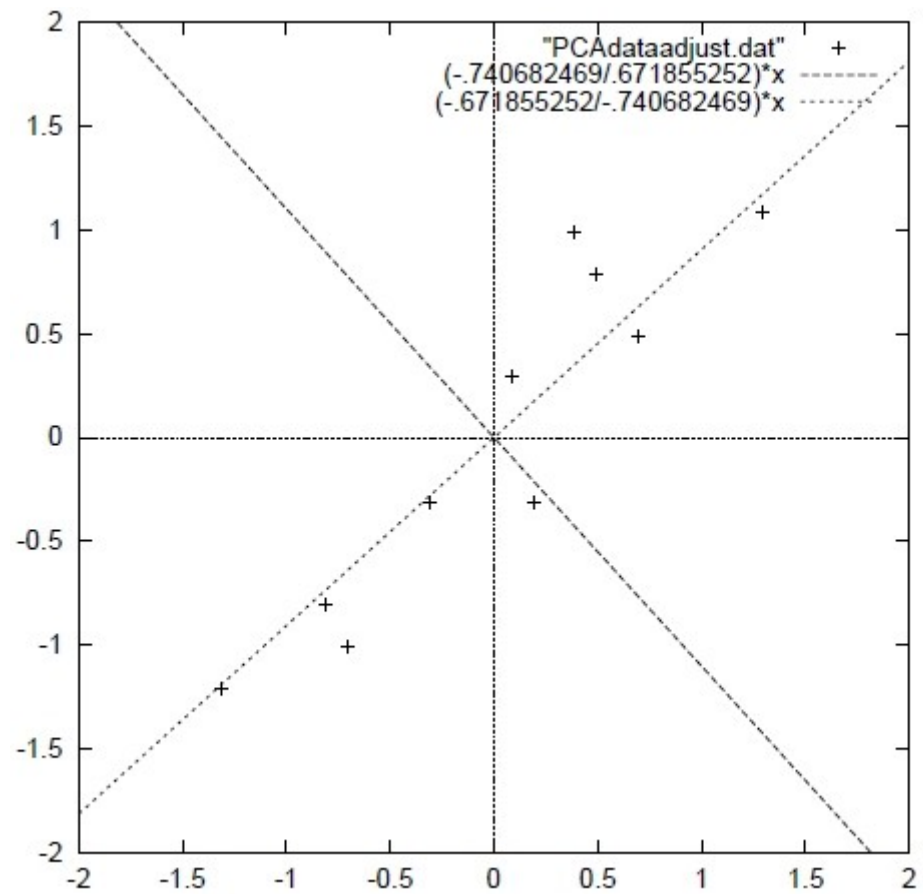
x	y		x	y
2.5	2.4		.69	.49
0.5	0.7		-1.31	-1.21
2.2	2.9		.39	.99
1.9	2.2		.09	.29
3.1	3.0	Normalized =	1.29	1.09
2.3	2.7		.49	.79
2	1.6		.19	-.31
1	1.1		-.81	-.81
1.5	1.6		-.31	-.31
1.1	0.9		-.71	-1.01

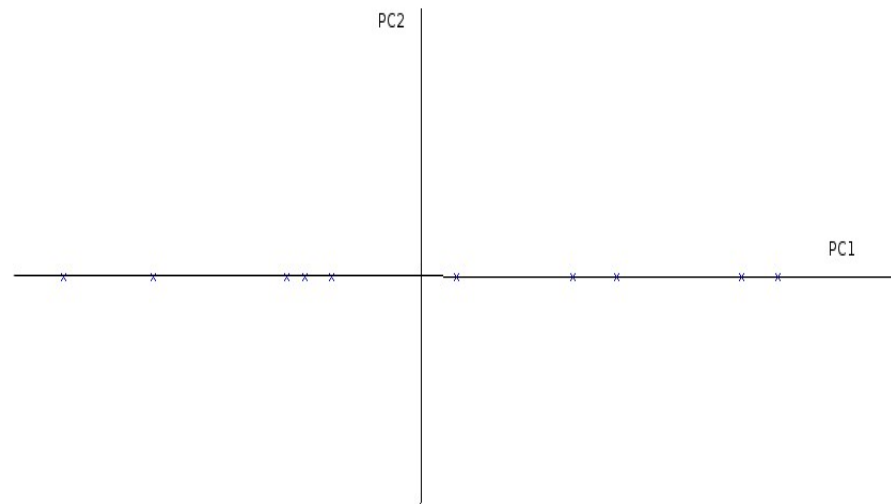
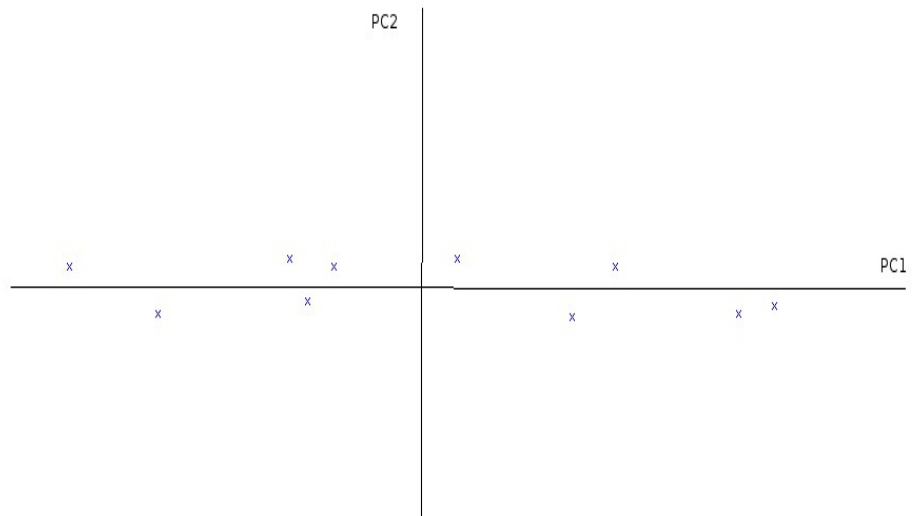


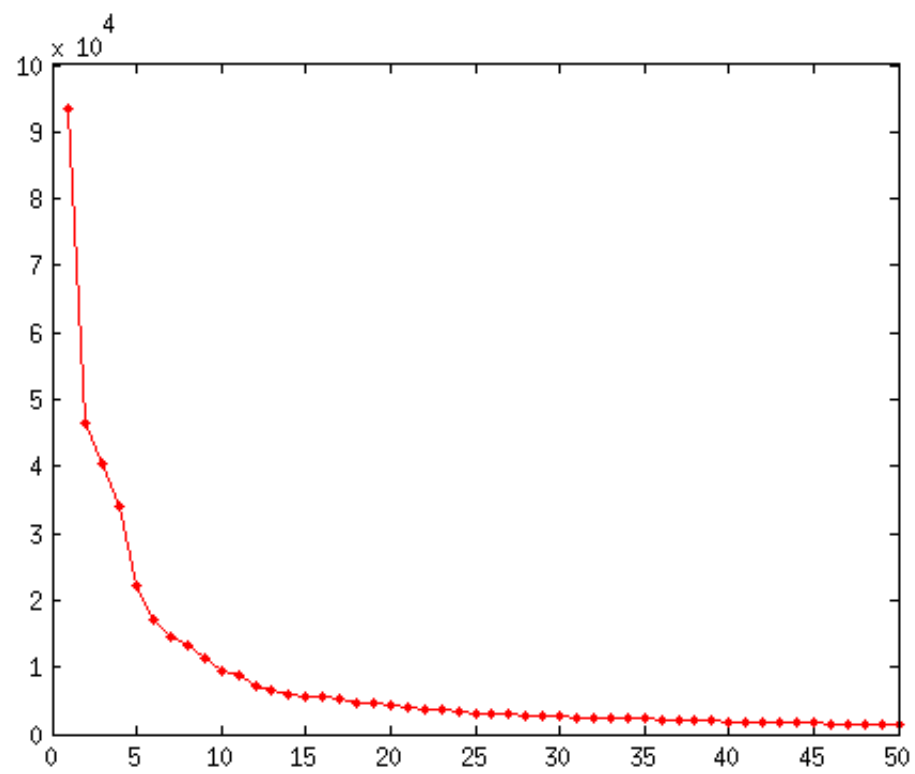
$$\text{Conv} = \begin{pmatrix} 0.6154444 & 0.7154444 \\ 0.6165555 & 0.6154444 \end{pmatrix}$$

$$\text{Eigenvalues} = \begin{pmatrix} 0.0490833989 \\ 1.28402771 \end{pmatrix}$$

$$\text{Eigenvector} = \begin{pmatrix} -0.735178656 & -0.677873399 \\ 0.677873399 & -0.735178656 \end{pmatrix}$$







Για την προβολή των δεδομένων στο χώρο των ιδιοδιανυσμάτων χρησιμοποιούμε τον τύπο:

$$w_i = v_i^T (X - \mu)$$

για $i=1..m$ όπου m η διάσταση του υπόχωρου

Για την ανακατασκευή των δεδομένων από τον αναμετασχηματισμένο χώρο:

$$X = Y + \mu$$

$$Y = \sum w_i v_i$$

όπου m η διάσταση του υπόχωρου

Dimensions = 120



Dimensions = 80



Dimensions = 160



Dimensions = 206

