ejemplo: para los datos de los billetes del banco suizo, el vector de medias

$$\overline{x}_{(n)} = \begin{pmatrix} 214.9 \\ 130.1 \\ 129.9 \\ 9.4 \\ 10.6 \\ 140.5 \end{pmatrix}$$

Los valores propios de $\hat{\Sigma}$, es decir, los elementos de la diagonal en la matriz $\mathcal{L}$ son

$$\begin{pmatrix} 2.985 \\ 0.931 \\ 0.242 \\ 0.194 \\ 0.085 \\ 0.035 \end{pmatrix}$$

Los vectores propios de $\hat{\Sigma}$, es decir las columnas $g_1, \dots, g_p$ de $\mathcal{G}$

$$\mathcal{G} = \begin{pmatrix} -0.0437 & 0.011 & -0.326 & 0.562 & 0.753 & 0.098 \\ 0.112 & 0.071 & -0.259 & 0.455 & -0.347 & -0.767 \\ 0.139 & 0.066 & -0.345 & 0.415 & -0.535 & 0.632 \\ 0.768 & -0.563 & -0.218 & -0.186 & 0.100 & -0.022 \\ 0.202 & 0.659 & -0.557 & -0.451 & 0.102 & -0.035 \\ -0.579 & -0.489 & -0.592 & -0.258 & -0.085 & -0.046 \end{pmatrix}$$

Como ejemplo calculamos las componentes principales para los datos de los billetes del banco suizo. La figura C nos muestra gráficos de dispersión para: (Y_1, Y_2) (Arriba a la izquierda), (Y_2, Y_3) (Arriba a la derecha), (Y_1, Y_3) (Abajo a la izquierda). Los caracteres "o" ó "+" corresponden a "verdaderos" ó "falsos". La gráfica abajo a la izquierda en la figura C, corresponde a los valores propios $l_1, l_2, \ldots, l_6$ de $\hat{\Sigma}$.

<u>Ejercicio</u>: Calcular de nuevo las componentes pero usando un reescalamiento de los datos $\tilde{\mathcal{X}}$. Por ejemplo si se asume que las variables X_1, X_2, X_3 y X_6 fueron medidos en cms. y que X_4 y X_5 se quedan como estaban originalmente, o sea en escala de mm., esto sería equivalente a re-escalar $\tilde{X}_1 = X_1/10$; $\tilde{X}_2 = X_2/10$; $\tilde{X}_3 = \frac{X_3}{10}$ y $\tilde{X}_6 = X_6/10$ y usar la matriz de datos $\tilde{\mathcal{X}}$

Este ejercicio tiene el fin de ilustrar que los gráficos pueden cambiar mucho. Dicho de otra forma: "LAS COMPONENTES PRINCIALES SON SENSIBLES A CAMBIOS DE ESCALA EN LAS VARIABLES".

Basta re-escalar una de las columnas en $\mathcal{X}$ para que los valores propios $l_1, \dots, l_p$ y los vectores propios $g_1, g_2, \dots, g_p$ cambien

Veremos más adelante que una forma de evitar este problema es trabajar con la matriz de correlaciones $\mathcal{R}$ de los datos $\mathcal{X}$.