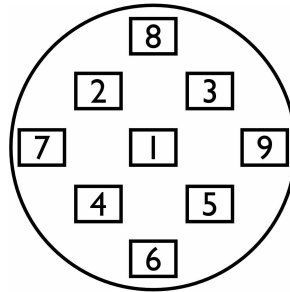


La solución a este taller debe subirse por SICUA antes de la 11:59PM del Lunes 02 de Octubre del 2017. Los archivos código fuente deben subirse en un único archivo `.zip` con el nombre `NombreApellido_hw2.zip`, por ejemplo yo debería subir el zip `VeronicaArias_hw2.zip`. Este archivo debe descomprimirse en un directorio de nombre `NombreApellido_hw2` que sólo contenga los códigos fuente `PCA.py` y `VelocidadCohete.py` (10 puntos). Recuerden que es un trabajo individual.

1. (40 points) **PCA**

Principal component analysis (PCA) es un método que permite reducir la dimensionalidad de un problema con muchas variables no independientes y que además permite explorar correlaciones entre dichas variables. La idea de este ejercicio es que usen los datos `siliconwaferthickness.csv` de grosor de wafers de silicona ([https://en.wikipedia.org/wiki/Wafer_\(electronics\)](https://en.wikipedia.org/wiki/Wafer_(electronics))). El grosor está medido en nueve puntos distintos del wafer como se muestra en la figura.



Para este ejercicio deben escribir un script de python (llamado `PCA.py`) con varias funciones que les permitan leer, organizar, hacer un análisis de componente principal y graficar los datos. El script debe:

- Leer los archivos de datos y guardar las variables relevantes en arrays.
- Normalizar los datos
- Graficar los datos y guardar la gráfica anterior (sin mostrarla) en `ExploracionDatos.pdf`
- Calcular la matriz de covarianza para los datos anteriores usando la implementación que hicieron individualmente en clase.
- Obtener e imprimir **en orden** en la consola los autovectores y autovalores obtenidos. Escriba también cuántas componentes principales considera que son necesarias para describir la variabilidad de los datos.
- Graficar los datos nuevamente en el sistema de referencia de los dos componentes principales. Esta gráfica debe ser clara, con ejes debidamente rotulados.
- Guardar la gráfica anterior (sin mostrarla) en `PCAdatos.pdf`
- Hacer una grafica donde se puedan ver las agrupaciones de las variables originales en el sistema de referencia de los dos componentes principales.

- Guardar la gráfica anterior (sin mostrarla) en `PCVariables.pdf`
- A partir de la gráfica anterior imprima las variables que están correlacionadas con `XX` donde `XX` corresponde a los grupos de variables encontrados.

2. (30 points) **Velocidad de un cohete**

Suponga que usted conoce las velocidades de un cohete para tres instantes de tiempo (ver tabla) y además usted sabe que la velocidad se comporta aproximadamente como un polinomio $v(t) = a_1 t^2 + a_2 t + a_3$. Usando los datos de la tabla y el método de eliminación gaussiana (implementado individualmente en clase) encuentre los coeficientes a_1 , a_2 y a_3 y encuentre la velocidad en $t = 7s$.

| Tiempo (s) | Velocidad (m/s) |
|------------|-----------------|
| 2 | 45.948 |
| 5 | 119.985 |
| 9 | 231.497 |

Para solucionar el problema debe escribir una rutina de Python llamada `VelocidadCohete.py` que:

- Plantee el sistema de ecuaciones descrito anteriormente en forma matricial.
- Solucione el sistema de ecuaciones usando los datos de la tabla y el método de eliminación gaussiana (implementado individualmente en clase) y encuentre los coeficientes a_1 , a_2 y a_3 .
- Haga una gráfica donde se muestren los valores de la tabla y el polinomio con los coeficientes encontrados por el método.
- Guardar la gráfica anterior (sin mostrarla) en `VelocidadCohete.pdf`
- Imprima en la consola los valores de a_1 , a_2 y a_3 obtenidos y el valor de la velocidad en $t = 7s$

Enlaces que les pueden ser útiles:

<http://webspace.ship.edu/pgmarr/Geo441/Lectures/Lec%2017%20-%20Principal%20Component%20Analysis.pdf>

<http://faculty.iiit.ac.in/~mkrishna/PrincipalComponents.pdf>