



**Benemérita
Universidad Autónoma de Puebla**

Facultad de Ciencias de la Computación

**“Análisis de Componentes
Principales utilizando los algoritmos
de Hessenberg y QR-Hessenberg”**

TESIS

que para obtener el grado de

**Licenciado en Ciencias de la
Computación**

PRESENTA:

Ángel Díaz Soto

Asesor:

M.C. Marcos González Flores

Puebla, Pue.

Diciembre 2009

Contenido

| | |
|--|-----------|
| INTRODUCCIÓN | 5 |
| CAPÍTULO 1. INGENIERÍA DE SOFTWARE | 8 |
| 1.1 ¿QUÉ ES LA INGENIERÍA DE SOFTWARE? | 8 |
| 1.2 MODELOS DE PRODUCCIÓN DE SOFTWARE | 8 |
| 1.3 MODELO EN CASCADA | 9 |
| 1.3.1 DEFINICIÓN Y ANÁLISIS DE REQUERIMIENTOS | 10 |
| 1.3.2 DISEÑO DEL SISTEMA | 11 |
| 1.3.3 CODIFICACIÓN Y PRUEBA DE LOS MÓDULOS | 11 |
| 1.3.4 INTEGRACIÓN Y PRUEBA DEL SISTEMA | 11 |
| 1.3.5 MANTENIMIENTO | 11 |
| 1.4 GENERALIDADES DE LOS MODELOS DE PRODUCCIÓN DE SOFTWARE | 12 |
| 1.4.1 FASE DE DEFINICIÓN | 12 |
| 1.4.2 FASE DE DESARROLLO | 13 |
| 1.4.3 FASE DE MANTENIMIENTO | 16 |
| CAPÍTULO 2. ANÁLISIS Y DISEÑO DEL SISTEMA | 18 |
| 2.1 ANÁLISIS DE REQUERIMIENTOS | 18 |
| 2.1.1 DATOS DE ENTRADA | 19 |
| 2.1.2 DATOS DE SALIDA | 19 |
| 2.2 DISEÑO DEL SOFTWARE | 20 |
| 2.2.1 ARCHIVO | 20 |
| 2.2.2 MOSTRAR | 21 |
| 2.2.3 COMPONENTES | 21 |
| 2.2.4 GRÁFICAS | 22 |
| 2.3 CONSIDERACIONES FINALES DEL DISEÑO | 22 |
| CAPÍTULO 3. DISEÑO DE MÓDULOS | 24 |
| 3.1 MÓDULO ARCHIVO | 24 |
| 3.1.1 NUEVO | 24 |
| 3.1.2 NÚMERO DE COLUMNAS | 25 |
| 3.1.3 LEER COLUMNA | 25 |
| 3.1.4 LEER MATRIZ | 26 |
| 3.1.5 GUARDAR MATRIZ | 27 |
| 3.1.6 EXPORTAR | 28 |
| 3.2 MÓDULO MOSTRAR | 29 |
| 3.2.1 MATRICES DE DATOS | 29 |

| | |
|--|------------------|
| 3.2.2 DATOS ESTÁNDAR | 29 |
| 3.2.3 DATOS CENTRADOS | 30 |
| 3.3 MÓDULO COMPONENTES | 30 |
| 3.3.1 COVARIANZA/DATOS ORIGINALES | 30 |
| 3.3.2 COVARIANZA/DATOS ESTÁNDAR | 31 |
| 3.3.3 COVARIANZA/DATOS CENTRADOS | 31 |
| 3.3.4 CORRELACIÓN/DATOS ORIGINALES | 32 |
| 3.3.5 CORRELACIÓN/DATOS ESTÁNDAR | 32 |
| 3.3.6 CORRELACIÓN/DATOS CENTRADOS | 33 |
| 3.4 MÓDULO GRÁFICAS | 33 |
| 3.4.1 DATOS ORIGINALES | 33 |
| 3.4.2 VALORES PROPIOS | 34 |
| 3.5 CODIFICACIÓN | 34 |
| 3.6 PANTALLA PRINCIPAL | 36 |
| | |
| <u>CAPÍTULO 4. ALGORITMOS DE HESSENBERG Y QR HESSENBERG</u> | <u>38</u> |
| | |
| 4.1 ALGORITMO DE HESSENBERG | 39 |
| 4.2 ALGORITMO DE QR-HESSENBERG | 39 |
| | |
| <u>CAPÍTULO 5. EJEMPLO DE EJECUCIÓN</u> | <u>43</u> |
| | |
| 5.1 DEFINIR DIMENSIONES | 43 |
| 5.2 LECTURA DE INFORMACIÓN | 45 |
| 5.2.1 LECTURA POR COLUMNA | 45 |
| 5.2.2 LECTURA POR MATRIZ | 47 |
| 5.3 GUARDAR MATRIZ DE DATOS | 49 |
| 5.4 EXPORTAR | 51 |
| 5.5 MOSTRAR MATRICES | 52 |
| 5.6 CÁLCULO DE COMPONENTES PRINCIPALES | 53 |
| 5.7 GRÁFICAS | 55 |
| 5.7.1 DATOS ORIGINALES | 56 |
| 5.7.2 VALORES PROPIOS | 57 |
| 5.8 BARRA DE HERRAMIENTAS | 57 |
| | |
| <u>CAPÍTULO 6. PRUEBAS</u> | <u>59</u> |
| | |
| 6.1 EJEMPLO 1 | 59 |
| 6.2 EJEMPLO 2 | 64 |
| | |
| <u>CAPÍTULO 7. CONCLUSIONES</u> | <u>69</u> |
| | |
| <u>ANEXO A. COMPONENTES PRINCIPALES</u> | <u>71</u> |

ANEXO B. CONSIDERACIONES SOBRE INSTALACIÓN Y EJECUCIÓN **72**

ANEXO C. DIAGRAMAS DE FLUJO **73**

ANEXO D. VISUAL C++ **75**

BIBLIOGRAFÍA **76**

Introducción

Los sistemas de software desde sus orígenes han ocupado un lugar importante en el área de la computación. En primer lugar, porque son una de las partes operacionales de una computadora, y en segundo lugar, por la gran cantidad de personas involucradas con ellos.

A principios de la era informática, el desarrollo de software se realizaba sin ninguna planificación y con gran esfuerzo, el hardware era de propósito general y el software se diseñaba a la medida de una solicitud ex profeso para resolver un problema; y por lo general una misma persona era quien lo escribía, lo ejecutaba, depuraba, y prácticamente no existía casi documentación. Hoy en día, existe una serie de disciplinas enfocadas a la labor de crear software conocida como “Ingeniería de Software” [1].

En esta tesis se aplican estas disciplinas para el desarrollo de un sistema enfocado a calcular las Componentes Principales de una matriz utilizando los algoritmos de Hessenberg y QR-Hessenberg [2].

Este sistema se diseñó para que sirviera como apoyo al laboratorio de espacios no invasivos de la BUAP, donde se realizan estudios de la superficie en el esmalte dental [3].

En el capítulo 1 se describe la ingeniería de software, sus características más importantes y los principales aspectos involucrados en el desarrollo de sistemas informáticos. Se enuncian los elementos del modelo de cascada o ciclo de vida clásico, ya que dicho modelo garantiza la reducción del trabajo

involucrado, debido a las retroalimentaciones a través de sus etapas establecidas.

En el capítulo 2 se aplicaron las características descritas en el capítulo anterior, para desarrollar el sistema que fue el objeto de esta tesis. Aquí se culminó con los diagramas de árbol de la estructura del sistema.

El capítulo 3 muestra el desarrollo de cada uno de los módulos creados en el capítulo 2, en particular, los diagramas de flujo utilizados para representar los diseños.

El capítulo 4 presenta los algoritmos revisados que fueron implementados en el desarrollo del sistema, en los que se ponen de manifiesto la parte fundamental, como son los cálculos de componentes principales.

El capítulo 5 está enfocado a describir un ejemplo de ejecución paso a paso utilizando los componentes básicos del sistema.

El capítulo 6 muestra algunos resultados obtenidos con la ejecución del sistema.

En el capítulo 7 se dan algunas conclusiones relacionadas con el trabajo realizado.

El anexo A incluye una breve historia sobre los componentes principales.

El anexo B incluye las consideraciones sobre la instalación y ejecución del sistema.

El anexo C describe los fundamentos de los diagramas de flujo.

El anexo D resume algunas características del lenguaje empleado, que fue Visual C++.

Capítulo 1. Ingeniería de Software

1.1 ¿Qué es la Ingeniería de Software?

La ingeniería de software puede definirse como “El establecimiento y uso de principios de ingeniería robustos, orientados a obtener software económico que sea fiable y funcione de manera eficiente sobre máquinas reales”.

La ingeniería de software incluye el diseño de programas, así como la documentación requerida para desarrollarlos. Es aplicable a proyectos pequeños o grandes, siendo la calidad del producto la primordial preocupación.

La calidad en los sistemas involucra características como:

- **Confiabilidad:** que desempeñe una función requerida adecuadamente.
- **Seguridad:** cubre los aspectos de confidencialidad y derechos de acceso.
- **Claridad:** debe ser escrito con claridad y ser fácil de entender.
- **Utilidad:** que satisfaga las necesidades del usuario.
- **Economía:** que sea de bajo costo en su desarrollo, mantenimiento y en su uso.

Quizás el factor más importante para un sistema es la utilidad, es decir, que el producto satisfaga las necesidades del usuario, a la vez que sea fácil de usar, sencillo de aprender y que cuente con buena presentación.

1.2 Modelos de producción de software

La ingeniería de software es utilizada para construir, entregar y hacer evolucionar un producto de software a partir de una idea inicial. Para esto, es necesario, entre otras cosas, la definición de un modelo para llevar a cabo el proceso de producción del software. Dicho modelo permitirá clasificar y

controlar las diferentes actividades para el desarrollo y mantenimiento del producto.

Existen diferentes modelos para llevar a cabo el desarrollo de software, algunos de estos son:

- Modelo en cascada o ciclo de vida del software.
- Modelo de prototipos.
- Modelo de programación exploratoria.
- Modelo orientado a la reutilización.
- Modelo de espiral.

En lo que se refiere a la presente tesis, se utilizó el modelo en cascada o ciclo de vida del software para el desarrollo del sistema, por lo que se analiza más a detalle este modelo.

1.3 Modelo en cascada

En este modelo, el proceso de desarrollo del software se ve como una serie de actividades sucesivas, donde se mantiene una relación entre ellas. Después de que se definen sus partes y requerimientos; se desarrolla y finaliza una actividad, ésta se aprueba para su implementación y se procede a realizar lo mismo con la siguiente. Cada etapa requiere información de entrada, desarrollo de procesos y presentación de sus resultados, todos ellos bien definidos. Se denomina modelo en cascada, porque los productos pasan de un nivel a otro con cierta suavidad (ver figura 1.1).

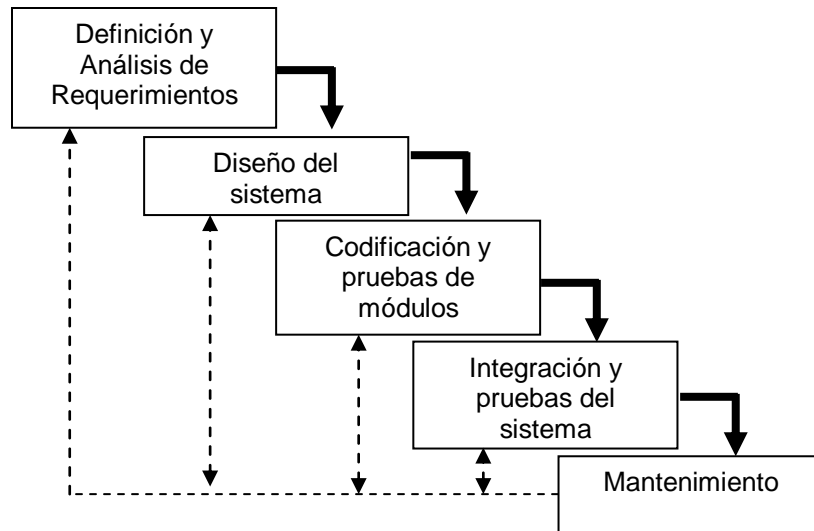


Figura 1.1 Modelo en cascada

En el ciclo de vida clásico:

- Cada proceso comienza cuando termina el anterior.
- Es difícil obtener la mayoría de requisitos al inicio.
- Es fácil de planificar.
- Es el ciclo ideal.

A continuación se explican las etapas de este modelo.

1.3.1 Definición y análisis de requerimientos

En esta etapa se establecen los servicios que satisfacen el sistema, sus restricciones y metas. Se establece una comunicación estrecha entre el cliente y el desarrollador, donde los dos participarán: uno interroga y el otro contesta, uno plantea sus necesidades y el otro actúa como quien dará solución de los problemas.

1.3.2 Diseño del sistema

El diseño es la descomposición del sistema en módulos, la descripción de lo que hace cada uno y sus interrelaciones. Los módulos son elementos que realizan funciones y subfunciones específicas.

1.3.3 Codificación y prueba de los módulos

En esta etapa se escriben los procedimientos para cada módulo y se prueba el correcto funcionamiento.

1.3.4 Integración y prueba del sistema

Cuando los módulos se prueban, son presentados para la integración del sistema. Conforme se avanza en la integración, es posible dar al cliente la oportunidad de ver el avance del producto, lo cual permite realizar más pruebas, y se pueden detectar puntos que hayan sido malentendidos, así como la interpretación de los mismos.

1.3.5 Mantenimiento

Una vez que el sistema sea instalado y puesto en funcionamiento, se realizarán las modificaciones para la corrección de errores que no se detectaron antes, o bien para integrar nuevas funciones a petición del cliente, por lo que el desarrollador debe involucrarse y tratar de anticiparse a dichas acciones.

El modelo en cascada no es el único disponible, pero si es el que más se utiliza, ya que las oportunidades para la retroalimentación dentro de él son muy grandes. Muchos desarrolladores ignoran todos los modelos de proceso de producción del software y proceden a generar el código directamente cuando

sólo una parte del diseño se ha concebido. Las presiones comerciales a menudo hacen esto casi inevitable, aunque se debe tener conciencia de que fallas en el proceso pueden dar lugar a resultados de un trabajo erróneo.

1.4 Generalidades de los modelos de producción de software

Aunque existen diferentes modelos de producción de software, todos tienen 3 fases en común (ver figura 1.2), las cuáles se explican a continuación.

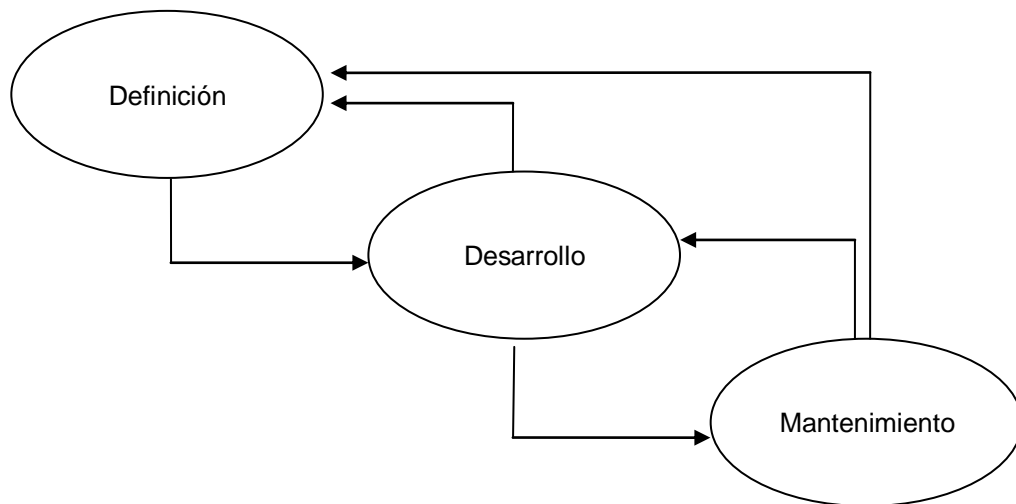


Figura 1.2. Fases genéricas del proceso de desarrollo

1.4.1 Fase de definición

Esta fase se enfoca sobre el qué; donde el desarrollador intenta identificar la información a ser procesada, la función y el rendimiento que se desea, etc.

Normalmente se divide en 3 pasos:

- Análisis del sistema: se define el papel de cada elemento del sistema.
- Planificación del proyecto de software: se asignan recursos, se estiman costos y se definen tareas.
- Análisis de requerimientos: se establece la función del sistema a través de las necesidades del usuario.

Aquí el desarrollador realizará una actividad intensa de comunicación con el cliente para que al final se cuente con los siguientes puntos:

- Dominio de la información y dominio funcional del problema: lograr el dominio de cómo los datos son procesados, transformados y cómo se generan los datos de salida. Para esto se podrán hacer entrevistas, cuestionarios e incluso se podría establecer el lugar donde se empleará el sistema, a fin de comprender totalmente cómo realizar el trabajo.
- Seccionamiento del problema: se tendrá que dividir el problema en partes entendibles y se establecerán interfaces entre ellas. Para esto se debe generar una representación jerárquica y después se dividirá cada uno de los elementos.
- Representación lógica del sistema: aquí se representan las funciones a realizar y la información a procesar. Para esto se podrá utilizar diagramas de flujo que se pueden ir refinando en niveles de mayor detalle.

1.4.2 Fase de desarrollo

Esta fase se enfoca al cómo; el que desarrolla intenta definir cómo han de diseñarse las estructuras de datos y cómo han de implantarse los procedimientos. También se divide en 3 pasos:

- Diseño del software: transforma los requerimientos del software en un conjunto de representaciones que describen las estructuras de datos y los procedimientos.
- Codificación: se trasladan las representaciones del diseño a algún lenguaje de programación.
- Prueba del software: se prueba el software para descubrir los defectos que pudieran existir.

Para lograr calidad, es de primordial importancia que el diseño sea modular, ya que se reduce la complejidad y produce una implementación más sencilla, incluso pudiendo permitir el desarrollo en paralelo de las diferentes partes del sistema. La fase de diseño puede dividirse en 4 puntos importantes:

- Diseño de datos: definir las estructuras de datos a utilizar.
- Diseño arquitectónico: define una estructura modular para el sistema.
- Diseño procedimental: transformar los módulos en una

descripción de los procedimientos del software.

- Diseño de la interfaz: mecanismos de interacción hombre-máquina.

La figura 1.3 resume los puntos anteriormente descritos.

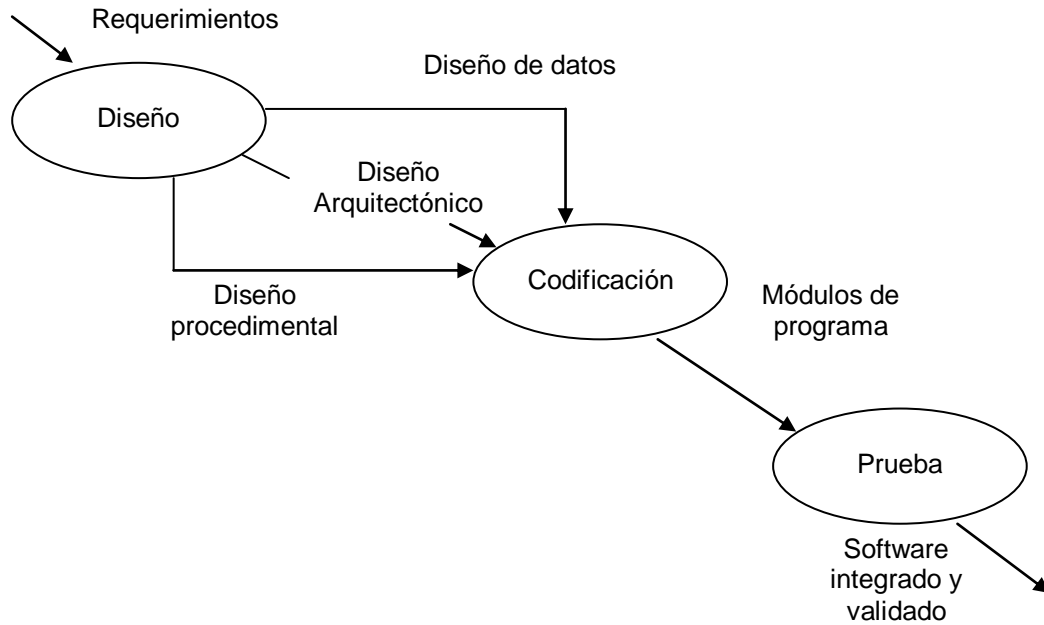


Figura 1.3 Fase de diseño

Para representar la estructura del sistema se pueden utilizar varias notaciones, pero la más común es un diagrama en forma de árbol, llamado "Diagrama de Estructura" que sirve para ver las relaciones entre los módulos (ver figura 1.4).

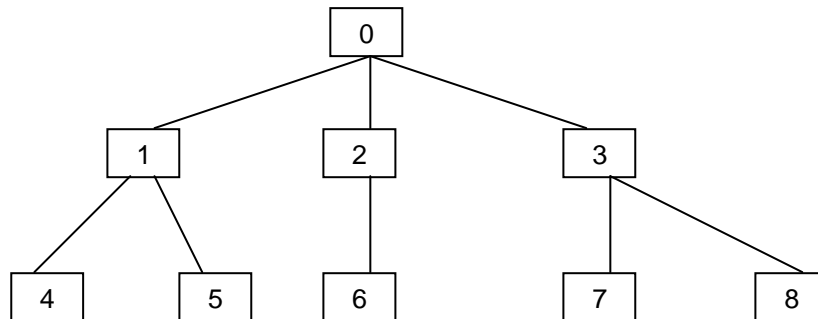


Figura 1.4 Diagrama de Estructura

Aquí es donde ya se muestra la estructura modular del sistema, que a su vez será dividida tantas veces como sea necesario, para que ayude a que un mejor entendimiento sea posible.

Posterior a la representación estructural del sistema, se deberá hacer la representación de los procedimientos. Como herramienta se puede utilizar alguna de las siguientes opciones:

- Árboles de decisión: diagrama en forma de árbol que presenta condiciones y acciones en forma secuencial (ver figura 1.5).

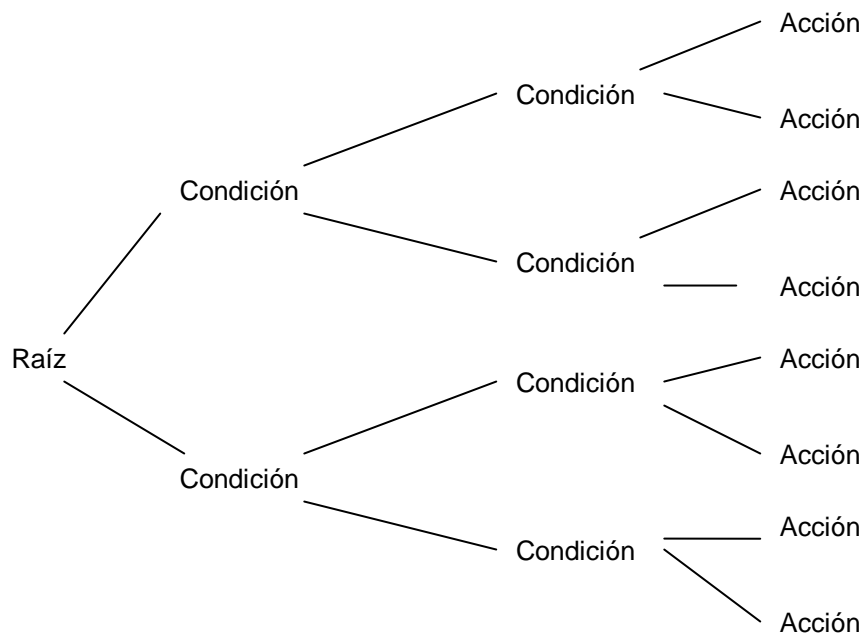


Figura 1.5 Árbol de decisión

- Tablas de decisión: matriz que muestra condiciones y acciones; las reglas de decisión se incluyen en la tabla y establecen qué procedimientos seguir, cuando existan ciertas condiciones (ver tabla 1.6).

| | |
|-------------|------------------------------|
| Condiciones | Combinaciones de condiciones |
| Acciones | Combinaciones de acciones |

Tabla 1.6 Tabla de decisión

- Diagramas de flujo: representación gráfica más utilizada, debido a su sencillez (ver anexo C). Fue la herramienta utilizada en el presente trabajo.
- Inglés estructurado: método empleado para superar los problemas del lenguaje ambiguo. Incluye tres tipos básicos de instrucciones (estructura secuencial, estructura de decisión y estructura de repetición).

1.4.3 Fase de mantenimiento

Esta fase se enfoca a la corrección de errores o adaptaciones. Esta fase re replica los pasos de las fases anteriores pero en el contexto del software ya existente. Aquí se encuentran 3 tipos de cambios:

- Corrección: se cambia el software para corregir errores.
- Adaptación: se traduce en modificaciones del software para acomodarlos a los cambios del entorno.
- Aumento: se incrementa el software por nuevos requerimientos del usuario.

En esta etapa, cuando exista alguna solicitud de cambio, se debe evaluar la relación costo/beneficio que pudiera implicar; si el cambio es de importancia considerable, debe ser notificado de inmediato. De ser posible la modificación, se harán los cambios necesarios produciendo una nueva versión del software, que igual que su predecesor, deberá ser puesto a prueba para verificar los

efectos del cambio. Por último, se deberá probar la operación del sistema como un todo para detectar cualquier efecto no deseado generado por el cambio.

Capítulo 2. Análisis y diseño del sistema

Una vez que se han repasado los fundamentos de la ingeniería de software, es el momento de aplicar sus pasos en el desarrollo del sistema. En particular, en este capítulo se utilizó el modelo en cascada haciendo primero un análisis de requerimientos y, posteriormente, se realizó la parte inicial del diseño del sistema.

2.1 Análisis de requerimientos

Recordemos que se desarrolló un sistema que ayuda a calcular los componentes principales de una matriz de covarianza o de correlación, en particular, se determinan los valores propios de una matriz. Partiendo de los requerimientos del laboratorio de espacios no invasivos, donde se utilizará el sistema a desarrollar, se detectaron las siguientes características de hardware (tabla 2.1) y de software (tabla 2.2) de los equipos en los que se ejecutaron pruebas y en si, el sistema.

| Dispositivo | Características mínimas | Características recomendadas |
|-------------|-------------------------|------------------------------|
| Procesador | Pentium II | Pentium IV |
| Memoria Ram | 128 Mb | 256 o 512 Mb |
| CD-ROM | Opcional | Opcional |
| Monitor | SVGA color | SVGA color |
| Disco duro | 10 MB | 40 o superior |
| Mouse | De 2 botones | Óptico o inalámbrico |

Tabla 2.1 Hardware

| Sistema operativo | Características mínimas | Características recomendadas |
|-------------------|-------------------------|------------------------------|
| Windows | 9x | 2000 o XP |

Tabla 2.2 Software

2.1.1 Datos de entrada

El sistema es capaz de leer datos a partir de archivos de texto ubicados en algún medio de almacenamiento. Estos datos están organizados en forma de columnas, por lo que cada archivo representa una columna de la matriz a procesar. Se deberán leer tantos archivos o columnas como sea necesario.

2.1.2 Datos de salida

Los resultados del procesamiento de la información se muestran en pantalla para su análisis. La salida será como texto en formato de matriz y opcionalmente se podrá almacenar en algún archivo en disco. Anexo a estas capacidades, se tendrán herramientas gráficas para analizar los resultados.

Como representación lógica del sistema de nivel 0 (conocida comúnmente como modelo fundamental) se presenta la siguiente figura (figura 2.1).

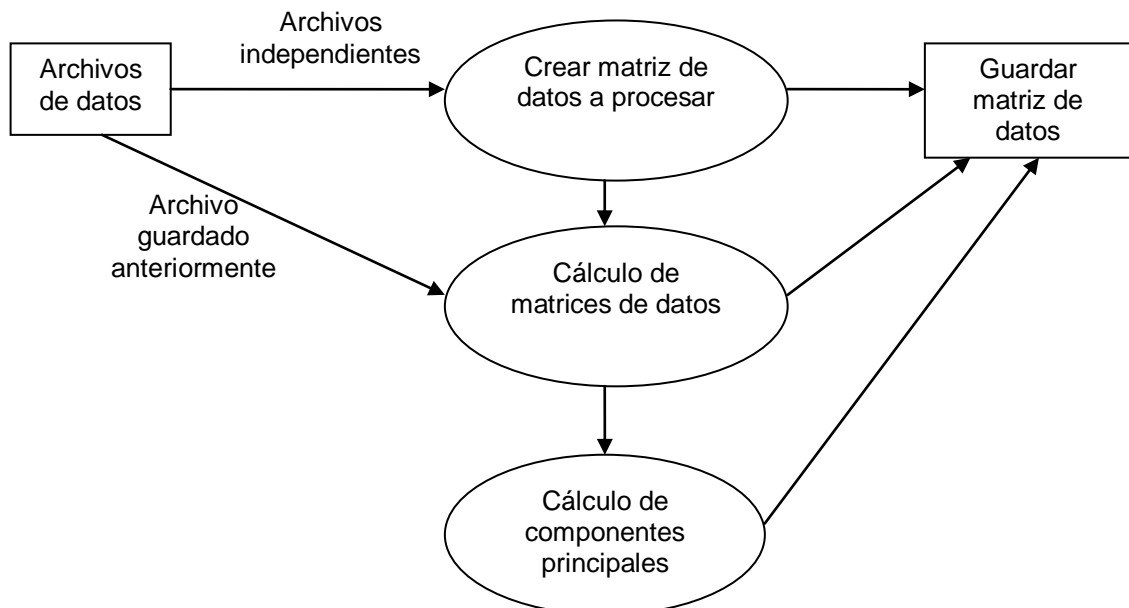


Figura 2.1 Representación lógica del sistema (diagrama de flujo de datos).

2.2 Diseño del software

Del análisis de la información obtenida del paso anterior (análisis de requerimientos) se obtiene el diagrama de estructura del sistema desarrollado (ver figura 2.2).

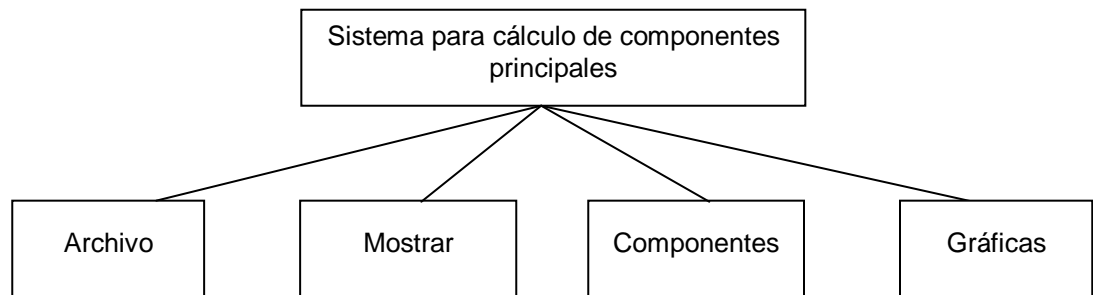


Figura 2.2 Diagrama de estructura del sistema

A continuación se detallan cada uno de los módulos que componen al sistema.

2.2.1 Archivo

El módulo archivo, para su operatividad, deberá mantener todas las características de los sistemas que se ejecutan en el ambiente de Windows, es decir, maneja opciones de abrir y guardar archivos, así como las de iniciar un proyecto nuevo, establecer las dimensiones de la matriz de datos, leer las columnas o una matriz completa (previamente guardada) y por último la opción de salir (ver figura 2.3)

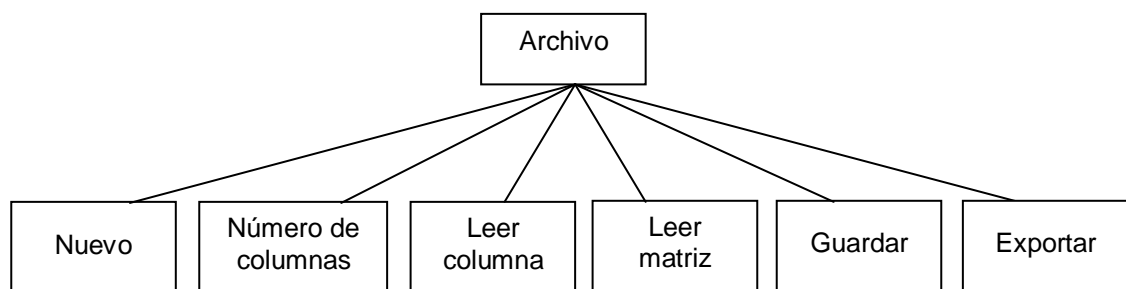


Figura 2.3 Módulo "Archivo"

2.2.2 Mostrar

El módulo mostrar mantiene las opciones de cálculo de matrices de datos básicas. Tales matrices de datos incluyen: la media, la varianza, la covarianza y la correlación. También se encuentran en este módulo las opciones Datos estándar y Datos centrados (vea figura 2.4).

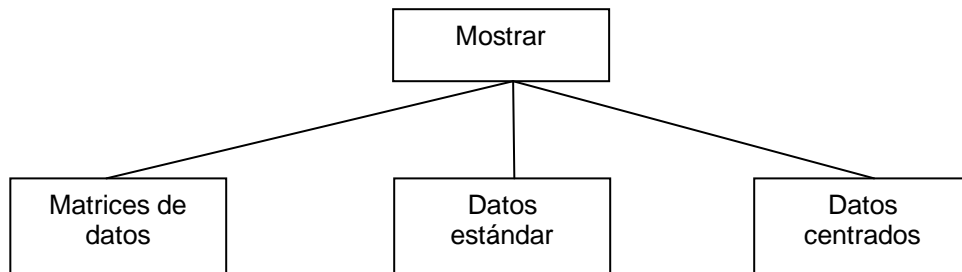


Figura 2.4 Módulo "Mostrar"

2.2.3 Componentes

Este módulo cuenta con las tareas específicas sobre el cálculo de componentes principales. En específico se divide en 2 partes: varianza y covarianza. Cada una de estas partes contará con 3 submódulos a su vez: datos originales, datos estándar y datos centrados (vea fig. 2.5).

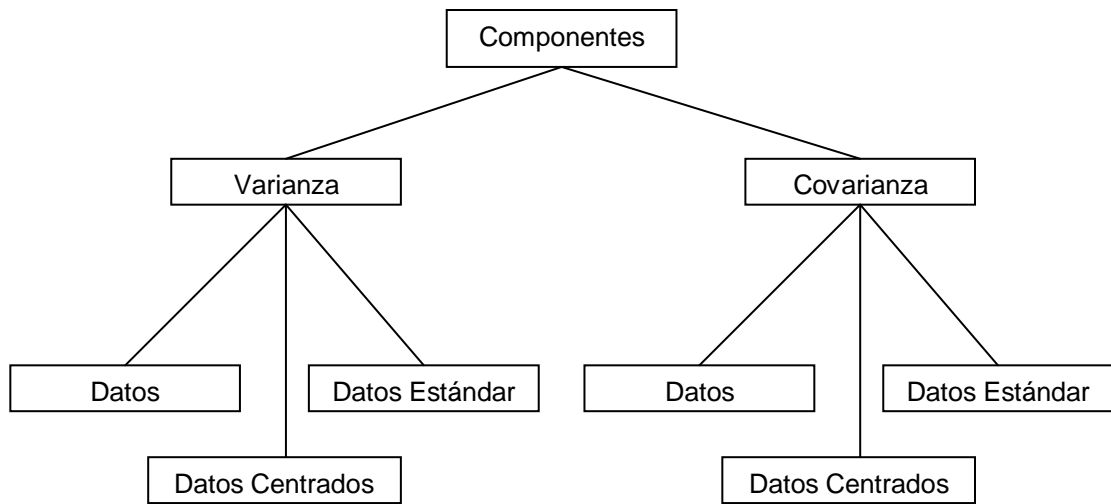


Figura 2.5 Módulo "Componentes"

Como puede verse, este módulo es el más grande de todos, ya que aquí es donde se hace empleo de los algoritmos de Hessenberg y QR-Hessenberg (vea el capítulo 4).

2.2.4 Gráficas

Este módulo permite ver algunos resultados en forma gráfica. En particular, se muestran los datos originales y los valores propios (ver figura 2.6). Cada módulo muestra en una ventana las gráficas solicitadas.

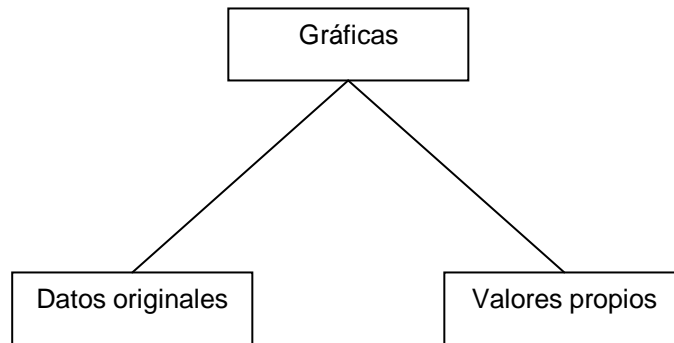


Figura 2.6 Módulo “Gráficas”

2.3 Consideraciones finales del diseño

Como parte importante del sistema, la información que se maneja para realizar las tareas encomendadas, debe estar disponible en todo momento para cualquier módulo del mismo. Para manejar la información, se utiliza una estructura que mantiene información sobre los datos originales, las matrices de datos principales (media, varianza, etc.) y las matrices con los resultados obtenidos, que a continuación se detallan:

- Matriz de datos originales.
- Vector de medias.
- Vector de varianzas.
- Matriz de covarianza.

- Matriz de correlación.
- Matriz de resultados.
- Matriz de componentes.
- Vector de valores propios.
- Vector de varianza total.

Cabe mencionar que para manejar matrices de gran tamaño, se demandarán más recursos del equipo (memoria y procesador) por lo que es indispensable contar con un equipo como el recomendado en la fase de análisis de requerimientos. En particular, en este trabajo se hizo empleo de una matriz de datos de tamaño máximo de 200 filas por 100 columnas. De requerirse una matriz de mayor tamaño se deberá hacer pruebas al funcionamiento del sistema en combinación con los recursos del equipo de desarrollo.

En lo que respecta al diseño de la interfaz hombre-máquina, se tomaron en cuenta los siguientes puntos importantes:

- Ofrecer una respuesta visual y auditiva.
- Verificar cualquier acción destructiva.
- Producir mensajes de error significativos.
- Usar ventanas para organizar los diferentes tipos de información.
- Minimizar el número de acciones para la entrada de datos que realiza el usuario.

Capítulo 3. Diseño de módulos

Una vez que se obtuvieron los diagramas de estructura de todos los módulos involucrados en el capítulo anterior, en este capítulo se procede a presentar los procedimientos correspondientes con diagramas de flujo. Se escogió esta técnica de diseño ya que permite describir mejor los pasos a seguir por cada procedimiento. En cada opción se da una breve descripción de lo que se realiza, así como de los datos de entrada y salida utilizados.

3.1 Módulo Archivo

3.1.1 Nuevo

Esta opción permitió establecer las variables que se manejan en el sistema a un valor o a un tamaño establecido de inicio (con opción de poderse modificar posteriormente).

- Entrada: Número de filas y columnas con valor 0.
- Salida: Matrices y vectores de trabajo con su tamaño predefinido y con sus valores puestos a 0.

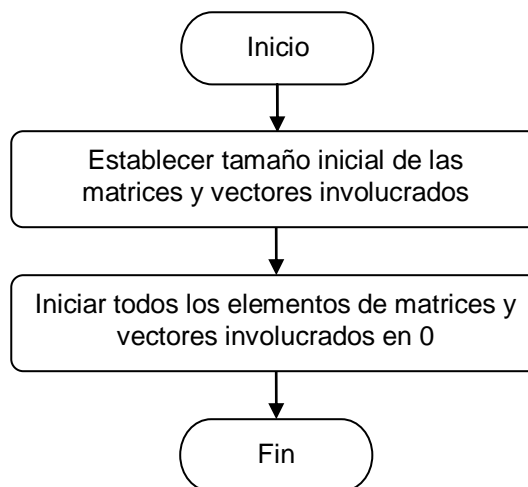


Figura 3.1 Nuevo

3.1.2 Número de columnas

Esta opción permite leer tanto el número de columnas como el número de filas de la matriz de datos originales desde el teclado. Se debe tomar en cuenta que el número de columnas, define también la longitud de algunos vectores, como la media y la varianza, así como el de valores propios.

- Entrada: Número de filas y columnas.
- Salida: Número de filas y columnas de la matriz de datos originales.

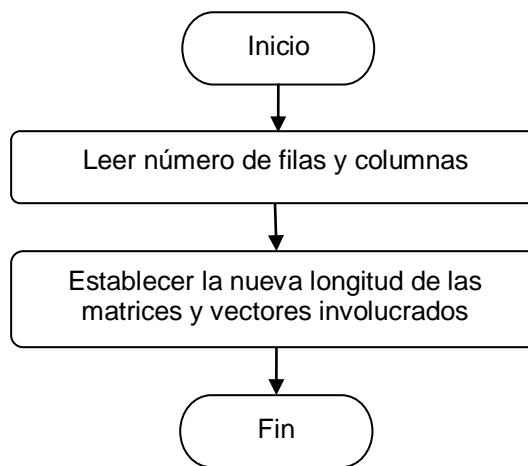


Figura 3.2 Número de columnas

3.1.3 Leer columna

Esta opción es la primera fuente de entrada de datos. Para esto se deberá seleccionar un archivo, que estará ubicado en algún medio de almacenamiento, y un número que representa la columna en la que se asignan los valores dentro de la matriz de datos.

- Entrada: Número de columna y nombre de archivo a leer.
- Salida: Valores de una columna seleccionada de la matriz de datos originales. Se cuenta con el posible cambio del número de filas, si el archivo contiene más datos y el usuario así lo desea.

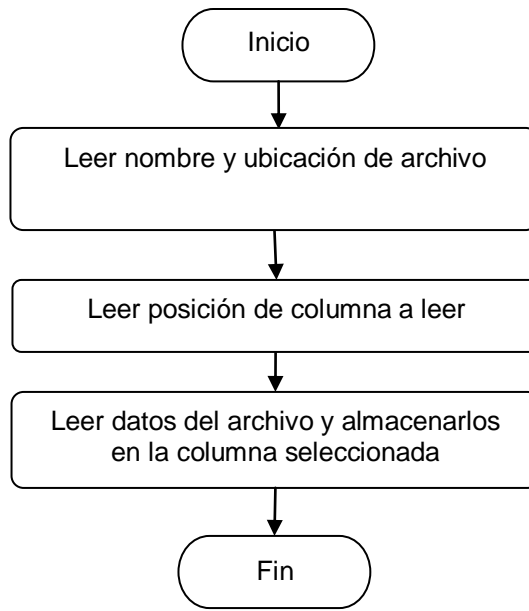


Figura 3.3 Leer columna

3.1.4 Leer matriz

Esta opción permitirá leer los datos desde algún archivo guardado previamente con la opción “Guardar matriz”. Si se intenta leer algún archivo que no haya sido generado con esa opción no podrá ser procesado. Al leer una matriz mediante esta opción se calculará automáticamente el vector de medias, el de varianzas, la matriz de covarianzas y la de correlación.

- Entrada: Nombre de archivo a leer.
- Salida: Valores de la matriz de datos originales, número de filas y columnas de la misma.

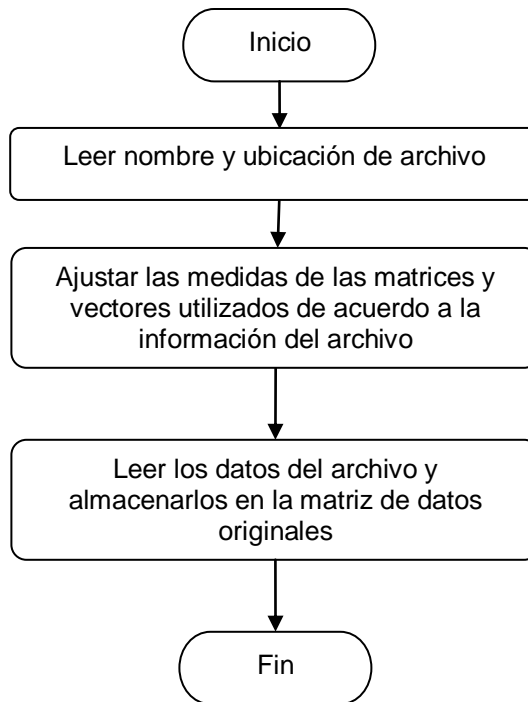


Figura 3.4 Leer matriz

3.1.5 Guardar matriz

Esta opción permite guardar los datos leídos en un archivo de texto para poder después leerse con la opción “Leer matriz”. El archivo resultante cuenta con una identificación que puede distinguirlo de cualquier otro tipo de archivo. Esta opción es muy útil para poder guardar los datos obtenidos al leer varias columnas.

- Entrada: Matriz de datos originales.
- Salida: Archivo con los valores de la matriz de datos originales e indicadores de número de filas y columnas de la misma.

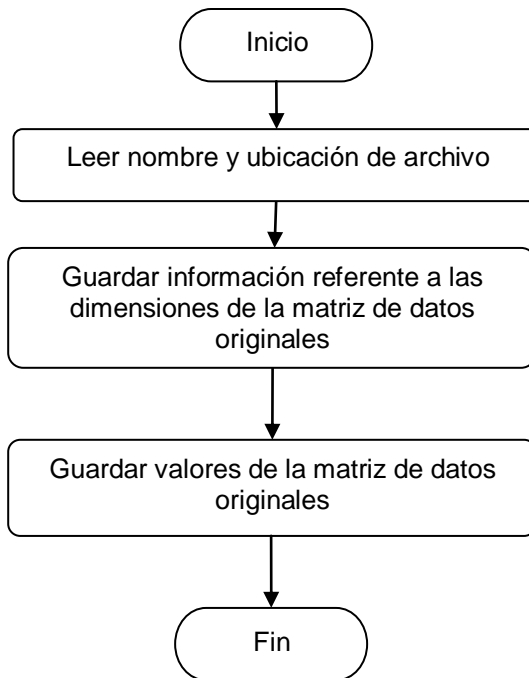


Figura 3.5 Guardar matriz

3.1.6 Exportar

Esta opción permite guardar la información de las matrices y vectores procesados en un archivo de texto.

- Entrada: Matriz de datos originales.
- Salida: Archivo con los valores de las matrices procesadas.

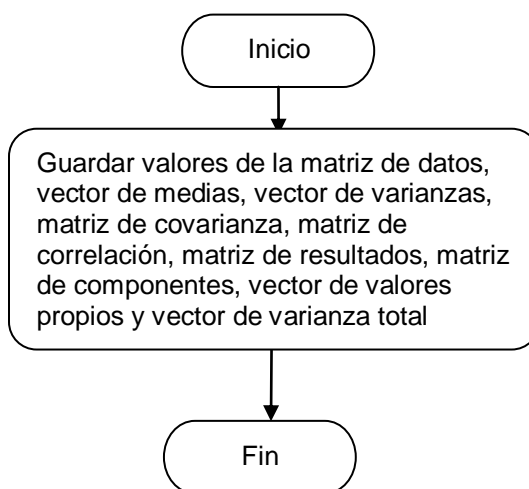


Figura 3.6 Exportar

3.2 Módulo Mostrar

3.2.1 Matrices de datos

Esta opción realiza el cálculo de la media, la varianza, la covarianza y la correlación. Los resultados que se muestran en pantalla.

- Entrada: Matriz de datos originales.
- Salida: Vector de medias, vector de varianzas, matriz de covarianza y matriz de correlación.

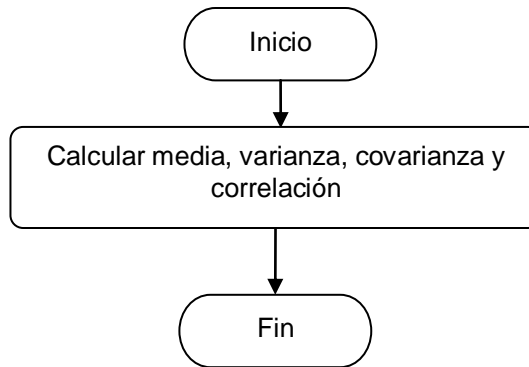


Figura 3.7 Matrices de datos

3.2.2 Datos estándar

Esta opción realiza el cálculo de la matriz con datos estandarizados.

- Entrada: Matriz de datos originales.
- Salida: Matriz de resultados con datos estandarizados.

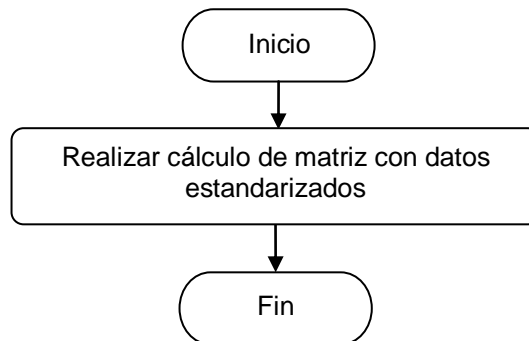


Figura 3.8 Datos estándar

3.2.3 Datos centrados

Esta opción realiza el cálculo de la matriz con datos centrados.

- Entrada: Matriz de datos originales.
- Salida: Matriz de resultados con datos centrados.

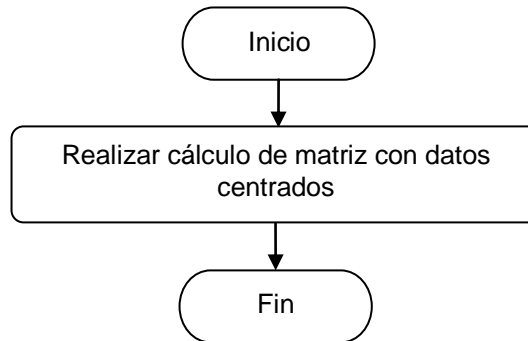


Figura 3.9 Datos centrados

3.3 Módulo Componentes

Este módulo se divide en 6 diagramas (muy similares) donde se realiza el cálculo estadístico de componentes principales. Cada diagrama se diferencia solamente por las matrices utilizadas. Debido a esto, solo se incluye en cada opción o procedimiento las indicaciones sobre entrada y salida utilizadas y el diagrama correspondiente.

3.3.1 Covarianza/Datos originales

- Entrada: Matriz de covarianza y matriz de datos originales
- Salida: Matriz de componentes, vector de valores propios y vector de varianza total.

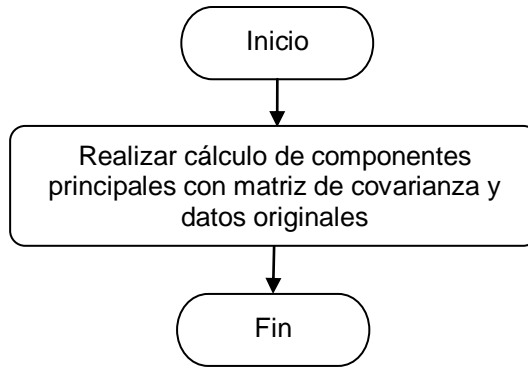


Figura 3.10 Covarianza/Datos originales

3.3.2 Covarianza/Datos estándar

- Entrada: Matriz de covarianza y matriz de resultados con datos estandarizados.
- Salida: Matriz de componentes, vector de valores propios y vector de varianza total.

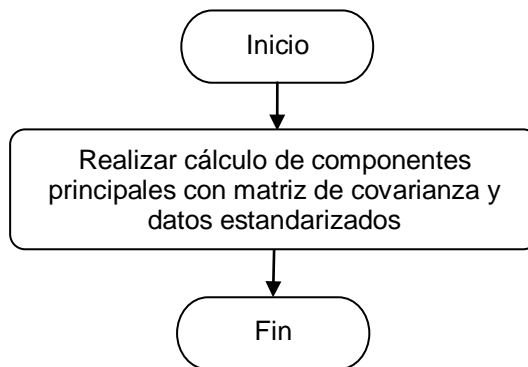


Figura 3.11 Covarianza/Datos estándar

3.3.3 Covarianza/Datos centrados

- Entrada: Matriz de covarianza y matriz de resultados con datos centrados.
- Salida: Matriz de componentes, vector de valores propios y vector de varianza total.

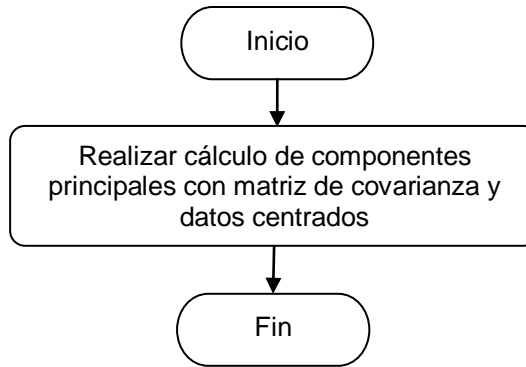


Figura 3.12 Covarianza/Datos centrados

3.3.4 Correlación/Datos originales

- Entrada: Matriz de correlación y matriz de datos originales.
- Salida: Matriz de componentes, vector de valores propios y vector de varianza total.

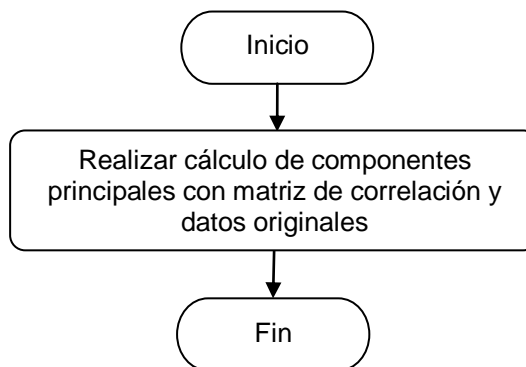


Figura 3.13 Correlación/Datos originales

3.3.5 Correlación/Datos estándar

- Entrada: Matriz de correlación y matriz de datos estandarizados.
- Salida: Matriz de componentes, vector de valores propios y vector de varianza total.

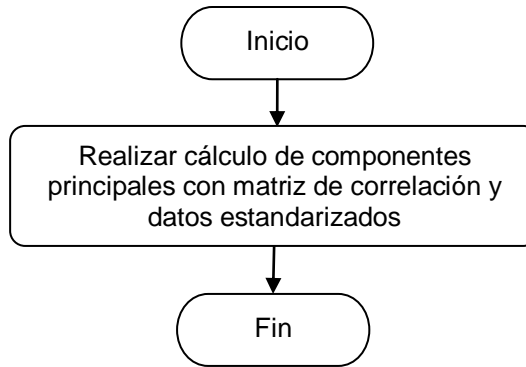


Figura 3.14 Correlación/Datos estándar

3.3.6 Correlación/Datos centrados

- Entrada: Matriz de correlación y matriz de datos centrados.
- Salida: Matriz de componentes, vector de valores propios y vector de varianza total.

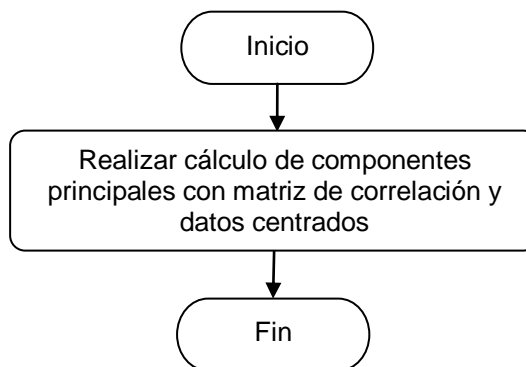


Figura 3.15 Correlación/Datos centrados

3.4 Módulo Gráficas

3.4.1 Datos originales

Esta opción muestra una gráfica con los valores de la matriz de datos originales. La gráfica tiene la capacidad de que el usuario decida qué columna graficar en cada momento, ya que en ocasiones no existe relación entre ellas y las escalas pudieran ser muy diferentes.

- Entrada: Matriz de datos originales.
- Salida: Ventana con la gráfica generada.

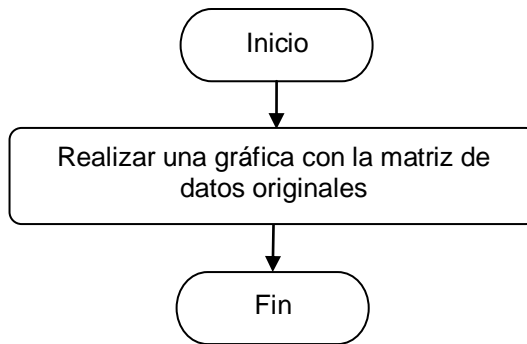


Figura 3.16 Gráficas/Datos originales

3.4.2 Valores propios

Esta opción muestra una gráfica con los valores propios obtenidos del cálculo de componentes principales.

- Entrada: Vector de valores propios.
- Salida: Ventana con la gráfica generada.

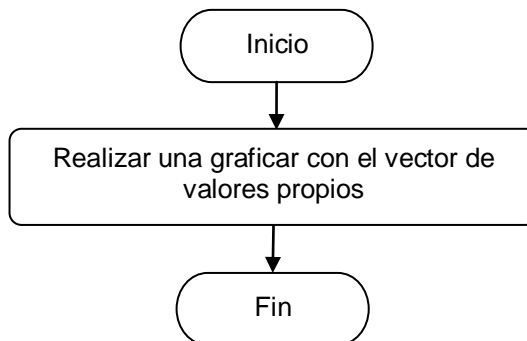


Figura 3.17 Gráficas/Valores propios

3.5 Codificación

Una vez terminado el diseño estructural del sistema y después de contar con las representaciones gráficas (diagramas de flujo) de los procedimientos a desarrollar, es tiempo de iniciar la codificación de cada uno de ellos. Durante la codificación, se trabajaron los módulos generados anteriormente como un menú del sistema desarrollado en Visual C++ [5], por lo que cada diagrama de flujo de un procedimiento se adecua a una opción del menú correspondiente.

Puesto que no existen reglas respecto al estilo de documentación a nivel de código fuente, sólo se puede seguir algunos principios generales como los mencionados a continuación:

- Utilizar sentencias GOTO de manera disciplinada.
- Utilizar sangrías, paréntesis, espacios y líneas en blanco.
- Evitar proposiciones “then” nulas.
- No anidar en forma muy profunda.
- Documentar adecuadamente el código.

Respecto a la documentación del código, es importante mencionar que se escogieron nombres significativos a los identificadores y que los comentarios descriptivos se emplearan para señalar lo que hacen bloques de código en lugar de comentar líneas.

En lo que se refiere a las sentencias de programación, éstas deben ser lo más simples posibles y seguir las siguientes recomendaciones:

- Evitar el uso de comparaciones condicionales complicadas.
- Eliminar las comparaciones con condiciones negativas.
- Usar paréntesis para mayor claridad en expresiones lógicas y aritméticas.
- Usar espacios para mayor legibilidad del contenido de las sentencias.

Respecto a los datos de entrada/salida, se debe seguir las recomendaciones siguientes:

- Validar todos los datos de entrada.
- Comprobar las combinaciones posibles de elementos de entrada.
- Usar indicativos de fin de datos en lugar de especificar número de elementos.
- Etiquetar las peticiones interactivas de entrada, especificando las posibles opciones o los valores límites.

3.6 Pantalla principal

Una vez que se terminan de codificar los procedimientos, el sistema queda listo para ser integrado. Incluso antes de ser terminado el sistema en su totalidad, ya se cuenta con lo que es la pantalla principal del sistema. El aspecto del sistema es amigable y muestra los resultados en la misma ventana. Esta ventana se muestra en la figura 3.18, mostrando los componentes principales de una ventana de Windows:

- Barra de título.
- Botones de minimizar, maximizar y cerrar.
- Barra de menús.
- Barra de herramientas.
- Área de trabajo.
- Barra de estado

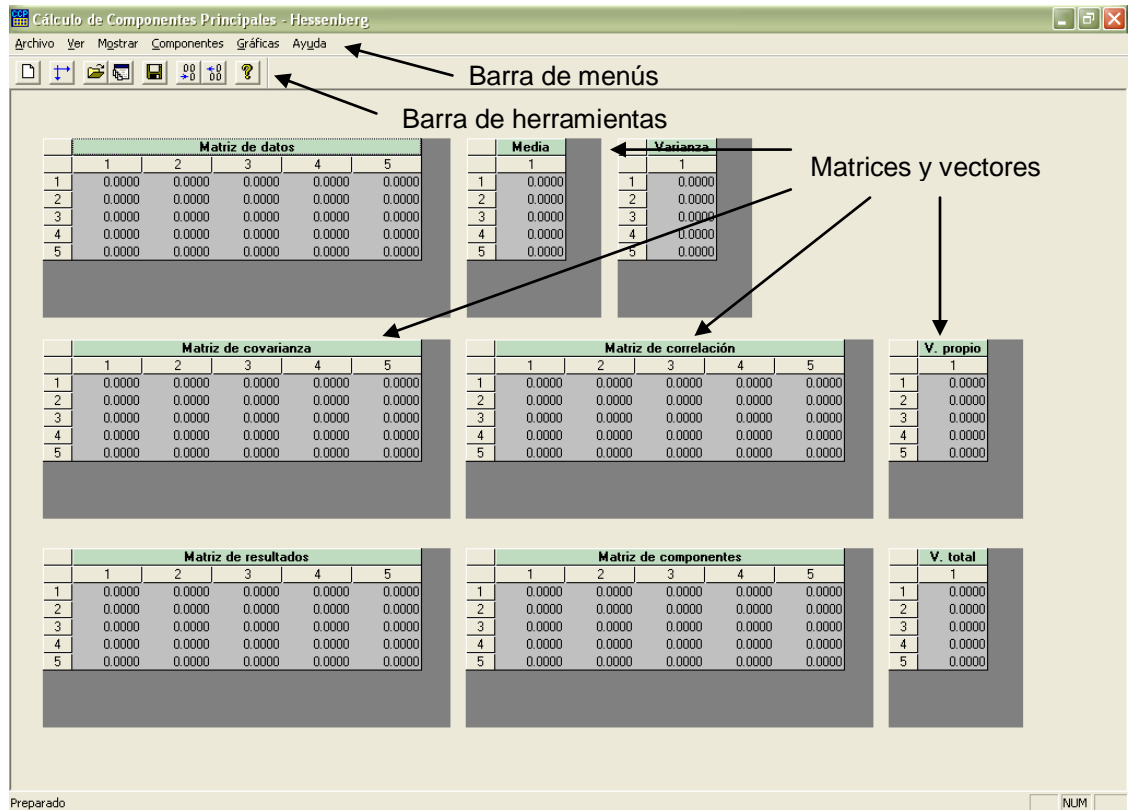


Figura 3.18 Pantalla principal del sistema

Como se ve en la figura anterior, en la barra de menús se encuentran los módulos principales (Archivo, Mostrar, Componentes y Gráficas).

También se encuentra disponible una barra de herramientas, con algunas opciones de ayuda para encontrar más fácilmente las opciones disponibles así como para ayudar a visualizar los datos mostrados (se explicarán más adelante).

Note también que en esta pantalla se muestran las matrices y vectores manejados por el sistema, en donde aparecerán los resultados de la lectura de archivos o de las tareas ejecutadas.

Capítulo 4. Algoritmos de Hessenberg y QR Hessenberg

En este capítulo se exponen los métodos utilizados para el cálculo de componentes principales de una matriz [6]. Para este cálculo existe la opción de usar la matriz de correlaciones o bien, la matriz de covarianzas. Cabe mencionar que los algoritmos ya se encuentran desarrollados y publicados [2], por lo que en este trabajo solamente se implementaron en el lenguaje seleccionado.

El método QR Hessenberg es un método numérico utilizado para calcular los valores y vectores propios de una matriz. El método consiste en utilizar el método QR aplicado a una matriz similar a la original pero del tipo Hessenberg.

No existe una forma sencilla de calcular los valores propios. Como el procedimiento involucra el cálculo del determinante de la matriz, se utiliza una gran cantidad de tiempo de procesamiento.

Una forma diferente para calcular los valores propios fue desarrollada a partir de 1950. La idea es crear una secuencia de matrices similares a la original, las cuales convergen a una matriz triangular superior. El algoritmo que produce esta secuencia de matrices usa la factorización QR.

Sin embargo, el método QR puede fallar y puede ser extremadamente lento. El algoritmo puede ser modificado para converger más rápidamente y para más matrices.

En la práctica el método QR es particularmente eficiente cuando la matriz es de tipo Hessenberg, ya que después de varias iteraciones sigue obteniendo matrices de este tipo y es más rápido que el método original.

4.1 Algoritmo de Hessenberg

```
Paso 1 Para m = 2,...,n-1
Paso 2 Tome x = 0.0
Paso 3 Tome i = m
Paso 4 Para j = m,...,n
Si  $|a_{j,m-1}| > |x|$  entonces
Tome  $x = a_{j,m-1}$ 
Tome i = j
Paso 5 Si  $i <> m$  entonces
  Para j=m-1,...,n
Tome  $y = a_{ij}$ 
Tome  $a_{ij} = a_{mj}$ 
Tome  $a_{mj} = y$ 
  Para j=1,...,n
Tome  $y = a_{ji}$ 
Tome  $a_{ji} = a_{jm}$ 
Tome  $a_{jm} = y$ 
Paso 6 Si  $x <> 0.0$  entonces
  Para i=m+1,...,n
Tome  $y = a_{i,m-1}$ 
Si  $y <> 0.0$  entonces
Tome  $y = y/x$ 
Tome  $a_{i,m-1} = y$ 
Para j=m,...,n
Tome  $a_{ij} = a_{ij} - y * a_{mj}$ 
Para j=1,...,n
Tome  $a_{jm} = a_{jm} + y * a_{ji}$ 
Paso 7 SALIDA
```

4.2 Algoritmo de QR-Hessenberg

```
Función signo(a,b)
Si  $b < 0.0$  entonces
Tome signo = -abs(a)
si no
Tome signo = abs(a)
SALIDA
```

```
Función min(a,b)
```

Si $a < b$ entonces
 Tome $\min = a$
 si no
 Tome $\min = b$
 SALIDA

Procedimiento QR_Hessenberg

Paso 1 Tome $\text{anorm} = |a_{11}|$
 Paso 2 Para $i=2, \dots, n$
 Para $j=i-1, \dots, n$
 Tome $\text{anorm} = \text{anorm} + |a_{ij}|$
 Paso 3 Tome $\text{nn} = n$
 Paso 4 Tome $t = 0.0$
 Paso 5 Mientras $\text{nn} >= 1$ Haga los pasos 6-10
 Paso 6 $\text{its} = 0$
 Paso 7 Para $l=\text{nn}, \dots, 2$
 Tome $s = |a_{l-1, l-1}| + |a_{ll}|$
 Si $s=0.0$ entonces
 Tome $s = \text{anorm}$
 Si $|a_{l, l-1}| + s = s$ entonces
 Ir a paso 9
 Paso 8 Tome $l = 1$
 Paso 9 Tome $x = a_{\text{nn}, \text{nn}}$
 Paso 10 Si $l=\text{nn}$ entonces
 Tome $w_{r_{\text{nn}}} = x+t$
 Tome $w_{i_{\text{nn}}} = 0.0$
 Tome $\text{nn} = \text{nn}-1$
 Si no
 Tome $y = a_{\text{nn}-1, \text{nn}-1}$
 Tome $w = a_{\text{nn}, \text{nn}-1} * a_{\text{nn}-1, \text{nn}}$
 Si $l=\text{nn}-1$ entonces
 Tome $p = 0.5 * (y-x)$
 Tome $q = p^2 + w$
 Tome $z = |q|^{1/2}$
 Tome $x = x+t$
 Si $q >= 0.0$ entonces
 Tome $z = p + \text{sign}(z, p)$
 Tome $w_{r_{\text{nn}}} = x+z$
 Tome $w_{r_{\text{nn}-1}} = w_{r_{\text{nn}}}$
 Si $z <> 0.0$ entonces
 Tome $w_{r_{\text{nn}}} = x - w/z$
 Tome $w_{i_{\text{nn}}} = 0.0$
 Tome $w_{i_{\text{nn}-1}} = 0.0$
 Si no
 Tome $w_{r_{\text{nn}}} = x+p$
 Tome $w_{r_{\text{nn}-1}} = w_{r_{\text{nn}}}$

Tome $w_{i_{nn}} = z$
Tome $w_{i_{nn-1}} = -z$
Tome $nn = nn-2$
Si no
Si $its = 30$ entonces
SALIDA('Número de iteraciones excedido')
(Procedimiento terminado sin éxito)
PARE
Si $its = 10$ o $its=20$ entonces
Tome $t = t*x$
Para $i=1, \dots, nn$
Tome $a_{ij} = a_{i^*}x$
Tome $s = |a_{nn,nn-1}| + |a_{nn-1,nn-2}|$
Tome $x = 0.75*s$
Tome $y = x$
Tome $w = -0.4375*(s)^{1/2}$
Tome $its = its+1$
Para $m=nn-2, \dots, 1$
Tome $z = a_{mm}$
Tome $r = x-z$
Tome $s = y-z$
Tome $p = (r*s-w)/a_{m+1,m} + a_{m,m+1}$
Tome $q = a_{m+1,m+1} - z - r - s$
Tome $r = a_{m+2,m+1}$
Tome $s = |p| + |q| + |r|$
Tome $p = p/s$
Tome $q = q/s$
Tome $r = r/s$
Si $m=1$ entonces
Ir a Etiqueta 4
Tome $u = |a_{m,m-1}| * (|q| + |r|)$
Tome $v = |p| * (|a_{m-1,m-1}| + |z| + |a_{m+1,m+1}|)$
Si $u+v = v$ entonces
 Ir a Etiqueta 4
Etiqueta 4: Para $i = m+2, \dots, nn$
Tome $a_{i,i-2} = 0.0$
Si $i <> m+2$ entonces
Tome $a_{i,i+3} = 0.0$
Para $k = m, \dots, nn-1$
Si $k <> m$ entonces
Tome $p = a_{k,k-1}$
Tome $q = a_{k+1,k-1}$
Si $k <> nn-1$ entonces
Tome $r = a_{k+2,k-1}$
Si no
Tome $r = 0.0$
Tome $x = |p| + |q| + |r|$

Si $x \neq 0.0$ entonces
 Tome $p = p/x$
 Tome $q = q/x$
 Tome $r = r/x$
 Tome $s = \text{sign}((p^2+q^2+r^2)^{1/2}, p)$
 Si $s \neq 0.0$ entonces
 Si $k=m$ then
 Si $l \neq m$ entonces
 Tome $a_{k,k-1} = -a_{k,k-1}$
 Si no
 Tome $a_{k,k-1} = -s*x$
 Tome $p = p+s$
 Tome $x = p/s$
 Tome $y = q/s$
 Tome $z = r/s$
 Tome $q = q/p$
 Tome $r = r/p$
 Para $j=k, \dots, nn$
 Tome $p = a_{kj}+q*a_{k+1,j}$
 Si $k \neq nn-1$ entonces
 Tome $p = p+r*a_{k+2,j}$
 Tome $a_{k+2,j} = a_{k+2,j}-p*z$
 Tome $a_{k+1,j} = a_{k+1,j}-p*y$
 Tome $a_{kj} = a_{kj}-p*x$
 Tome $mmin = \min(nn, k+3)$
 Para $i=1, \dots, mmin$
 Tome $p = x*a_{ik}+y*a_{i,k+1}$
 Si $k \neq nn-1$ entonces
 Tome $p = p+z*a_{i,k+2}$
 Tome $a_{i,k+2} = a_{i,k+2}-p*r$
 Tome $a_{i,k+1} = a_{i,k+1}-p*q$
 Tome $a_{ik} = a_{ik}-p$
 Ir Paso 7
 Paso 11 SALIDA

Capítulo 5. Ejemplo de ejecución

En este capítulo se muestra, mediante captura de pantallas para un ejemplo real, lo que representa una corrida del sistema desarrollado, y se indican los pasos a seguir para las siguientes actividades:

- Definir dimensiones de la matriz de datos.
- Leer la información a procesar.
- Guardar los datos leídos.
- Exportar.
- Ejecutar las tareas del sistema.
- Mostrar las gráficas disponibles.

5.1 Definir dimensiones

Este puede ser el primer paso a realizar antes de trabajar, aunque puede realizarse en cualquier momento. Para hacerlo, seleccione el menú “Archivo” y después la opción “Numero de columnas...” (vea la figura 5.1).

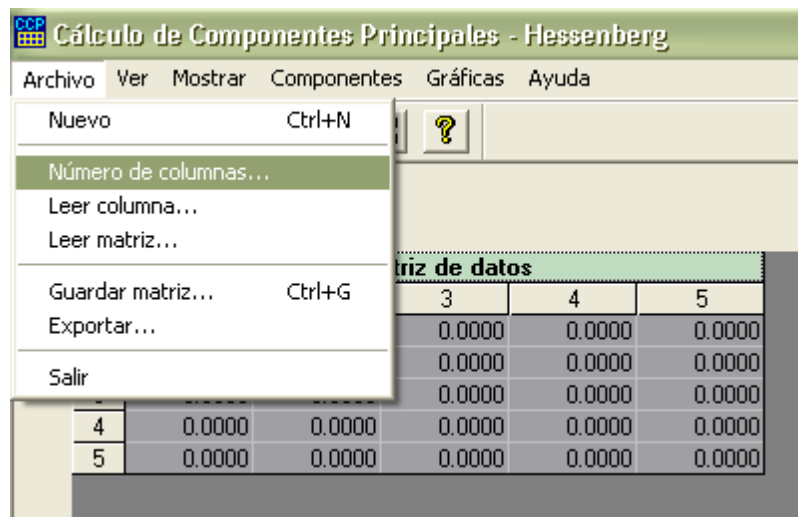


Figura 5.1

Al hacer lo anterior se mostrará una ventana para poder introducir los datos del número de filas y columnas (vea la figura 5.2)

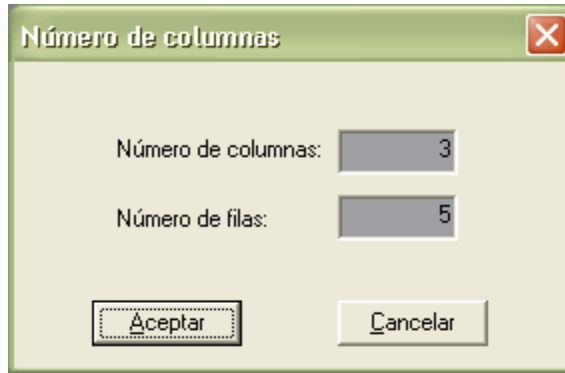


Figura 5.2

Se teclean los números deseados y se pide hacer clic en al botón “Aceptar”. Esto hará que las matrices y vectores mostrados en la ventana principal cambien a las medidas indicadas aquí. Vea la figura siguiente (figura 5.3) con un ejemplo de 3 columnas por 5 filas.

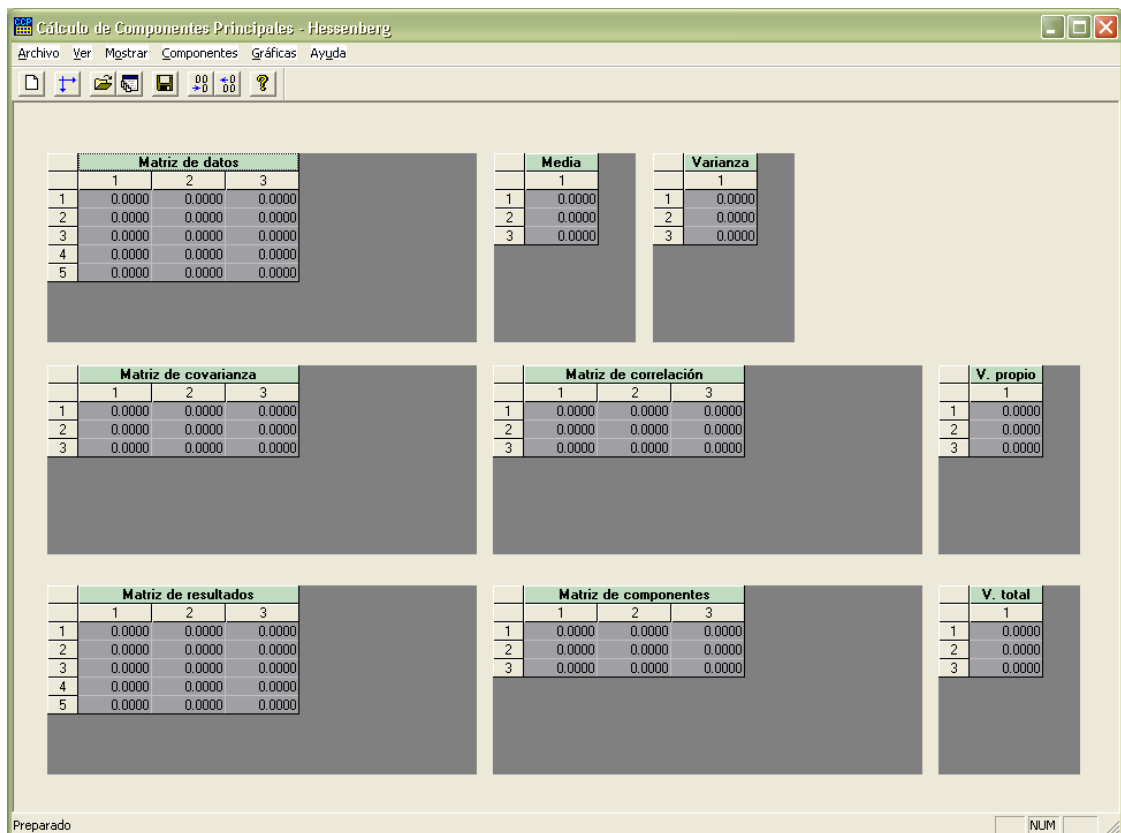


Figura 5.3

5.2 Lectura de información

La lectura de información permitirá que el sistema obtenga los datos para poder ejecutar las demás tareas. La lectura de información se puede realizar por 2 formas:

- Por columna
- Por matriz

5.2.1 Lectura por columna

Para realizar esta tarea seleccione el menú Archivo y después la opción “Leer columna...” como se muestra en la figura 5.4.

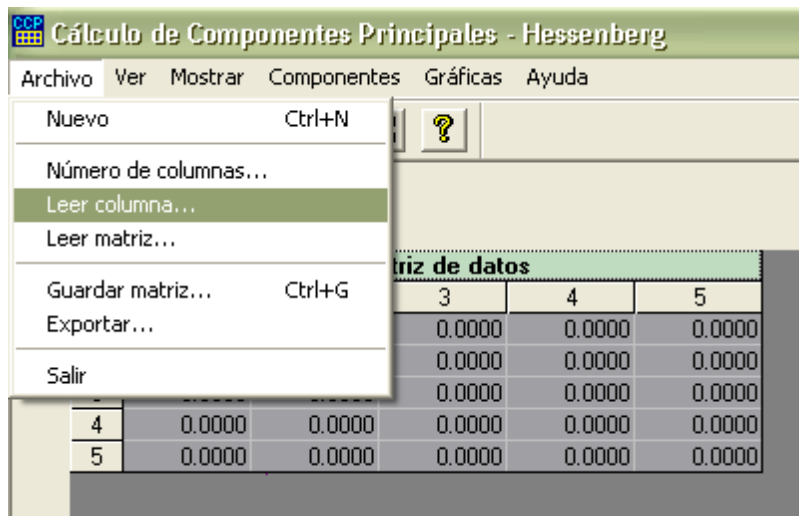


Figura 5.4

Aparecerá entonces la siguiente ventana (figura 5.5).

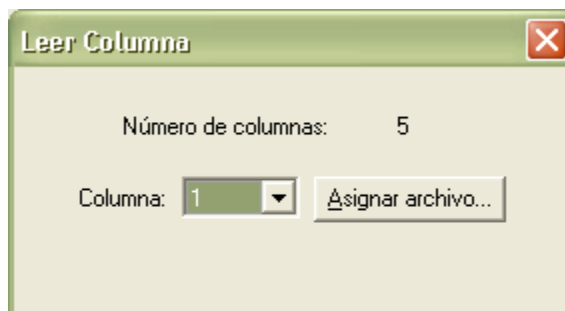


Figura 5.5

La lista columna, indica en cuál de éstas se va a introducir los datos contenidos en el archivo que se seleccione. Note que solo podrá seleccionar un número de columna de acuerdo a las dimensiones definidas anteriormente (vea figura 5.6).

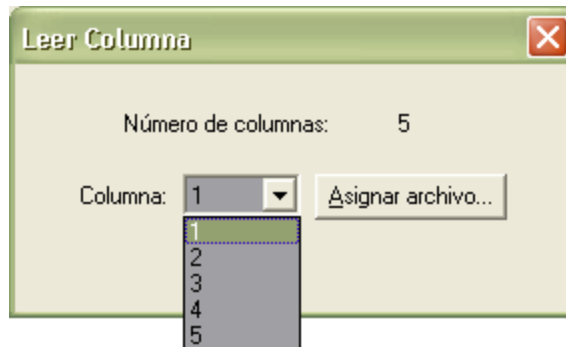


Figura 5.6

Para asignar el archivo deberá hacer clic en el botón “Asignar archivo...”. Al hacerlo aparecerá la ventana donde se podrá seleccionar el archivo (ver figura 5.7).

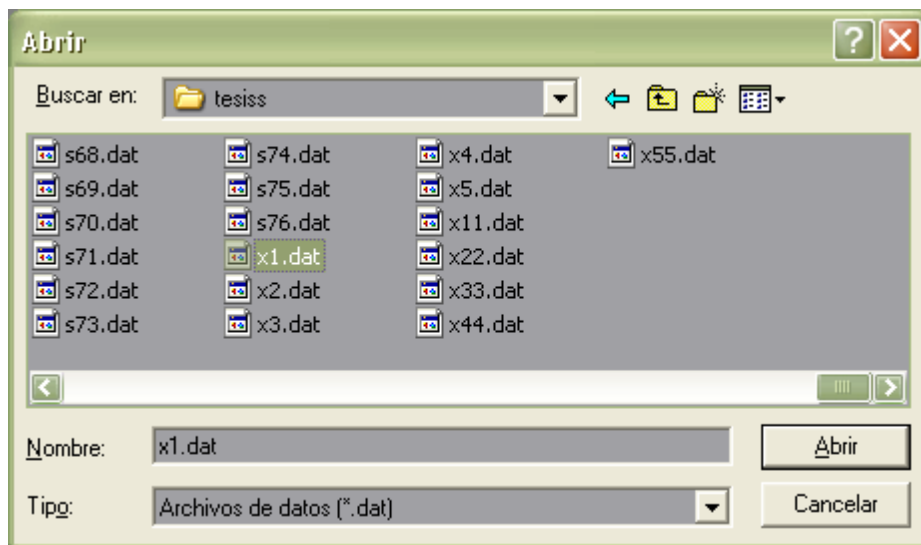


Figura 5.7

Una vez seleccionado el archivo y al hacer clic en el botón “Abrir”, aparecerá nuevamente la pantalla principal del sistema con los datos ya leídos mostrados

en la matriz de datos (vea la figura 5.8). Este proceso de leer columna lo deberá ejecutar tantas veces como archivos desee leer.

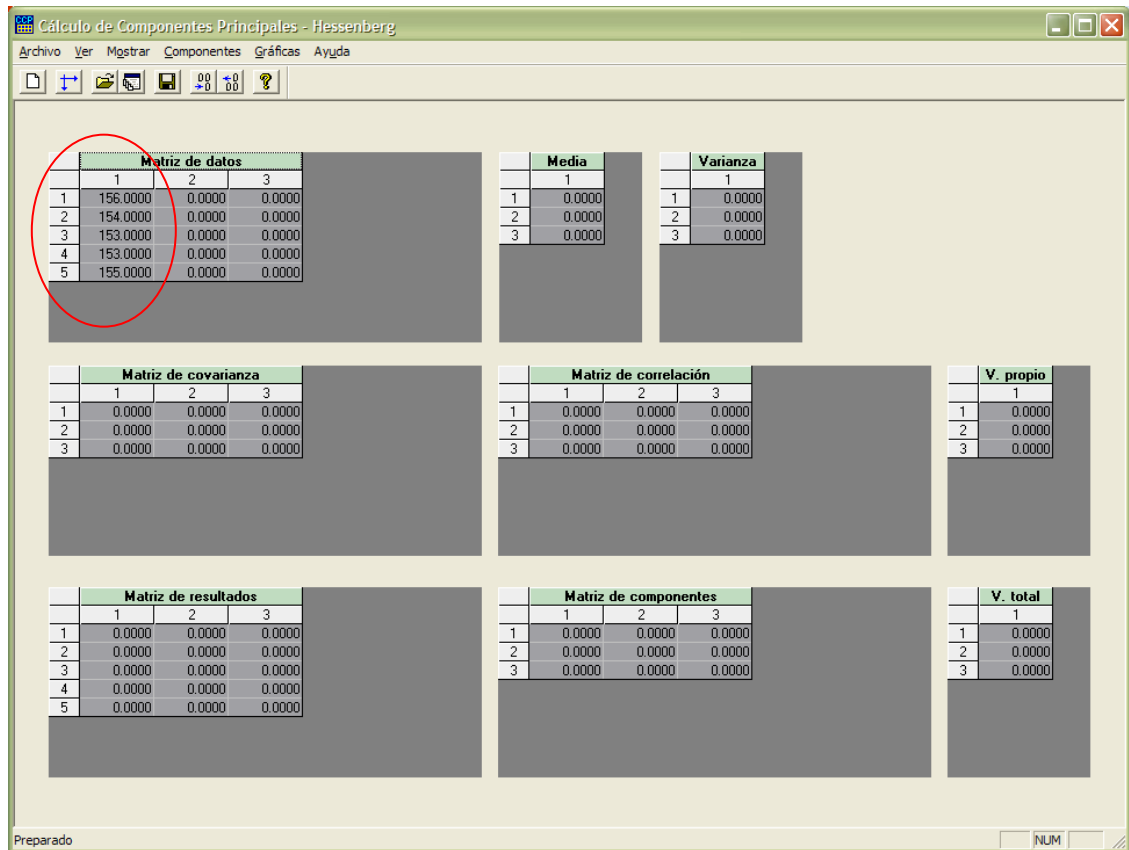


Figura 5.8

5.2.2 Lectura por matriz

Para realizar esta tarea seleccione el menú Archivo y después la opción “Leer matriz...” como se muestra en la figura 5.9.

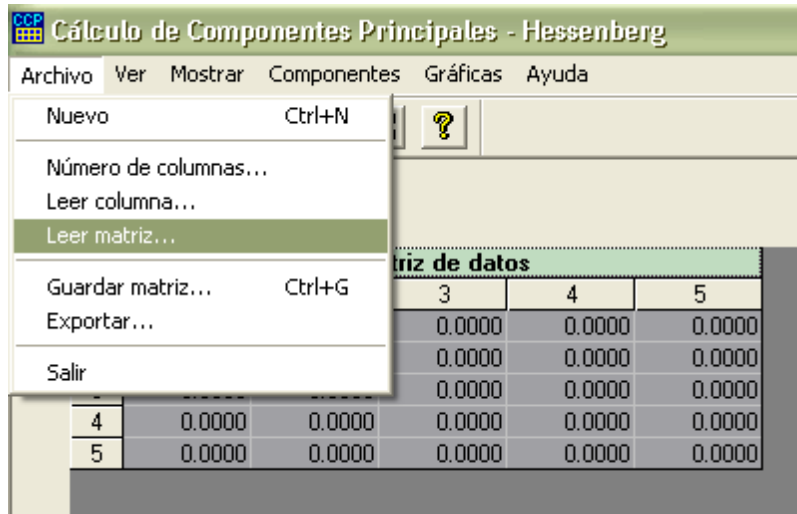


Figura 5.9

Deberá aparecer la siguiente ventana (figura 5.10):

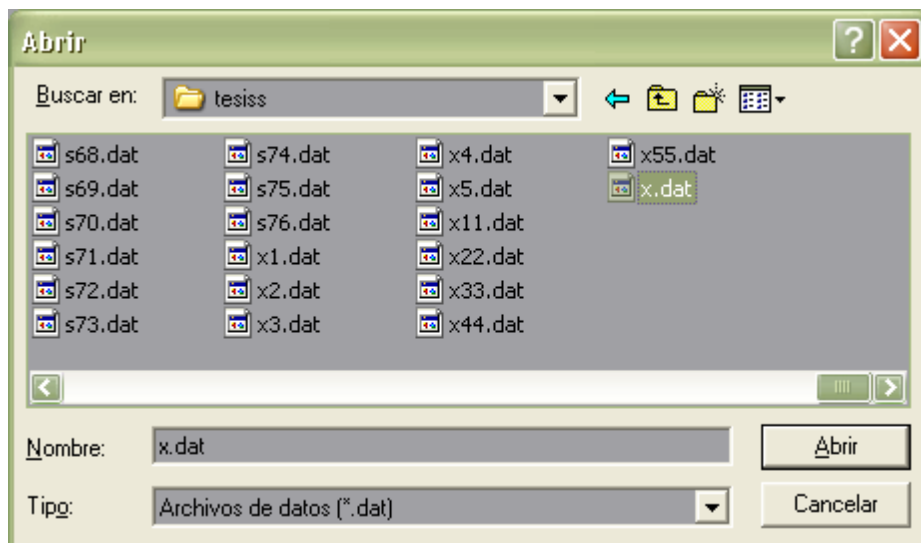


Figura 5.10

Aquí podrá seleccionar el archivo que previamente fue guardado con la opción “Guardar matriz...” (descrita más adelante). Al seleccionar el archivo y hacer clic en el botón “Abrir”, nuevamente regresará a la pantalla principal con los datos ya leídos (vea figura 5.11). Note que inmediatamente se calculan y se muestran los vectores de medias, de varianzas y las matrices de covarianza y de correlación.

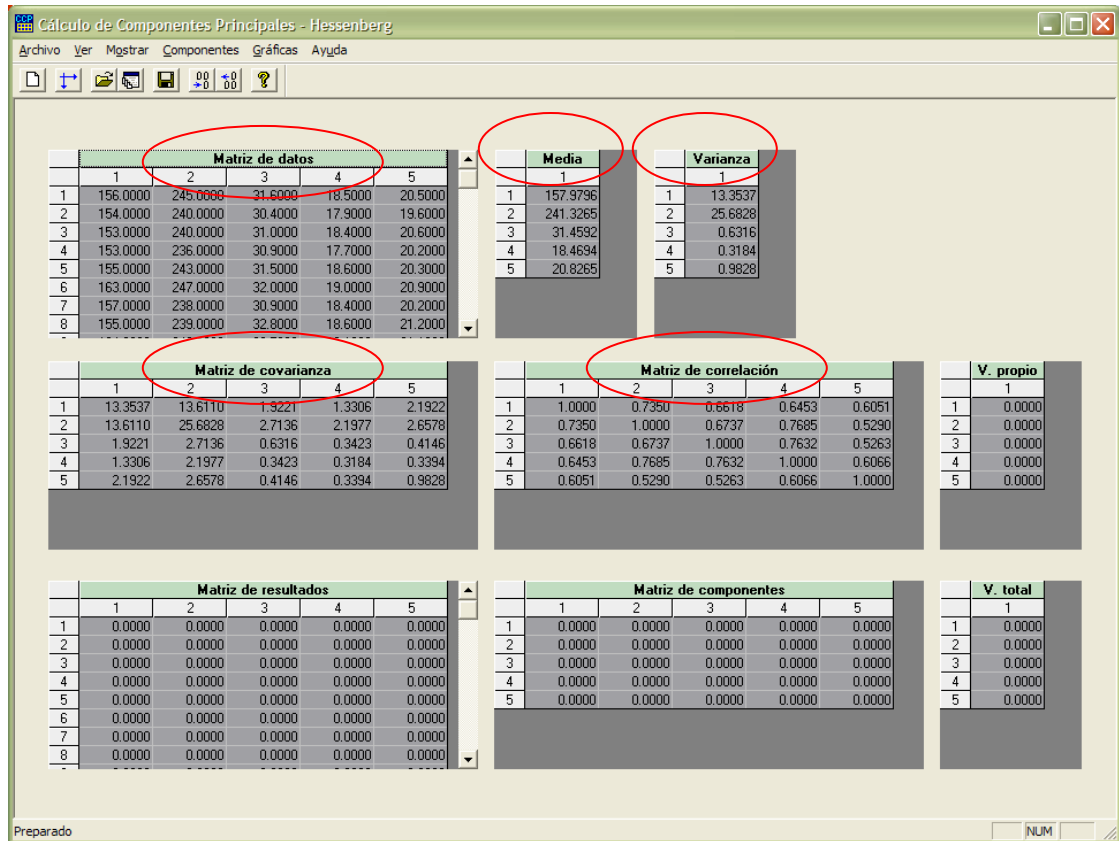


Figura 5.11

5.3 Guardar matriz de datos

Para guardar la información contenida en la matriz de datos, así como las dimensiones de esta matriz, seleccione la opción “Guardar matriz...” del menú “Archivo” como se muestra en la figura 5.12. Después de seleccionarla aparecerá la ventana mostrada en la figura 5.13.

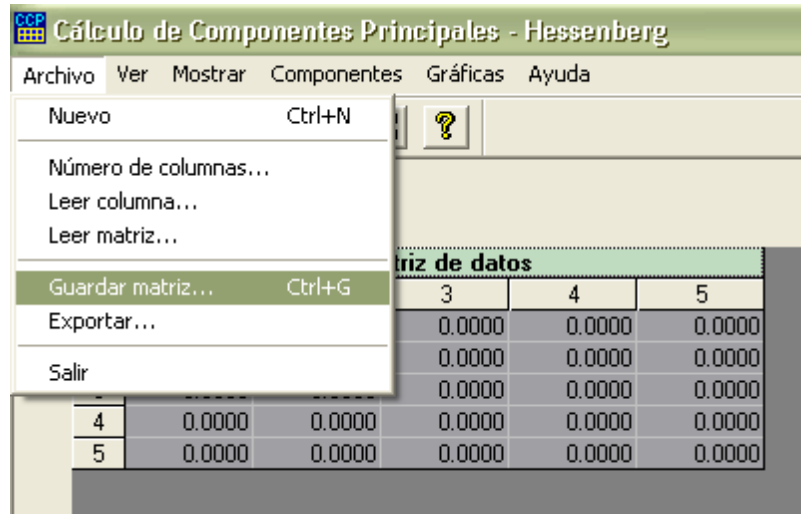


Figura 5.12

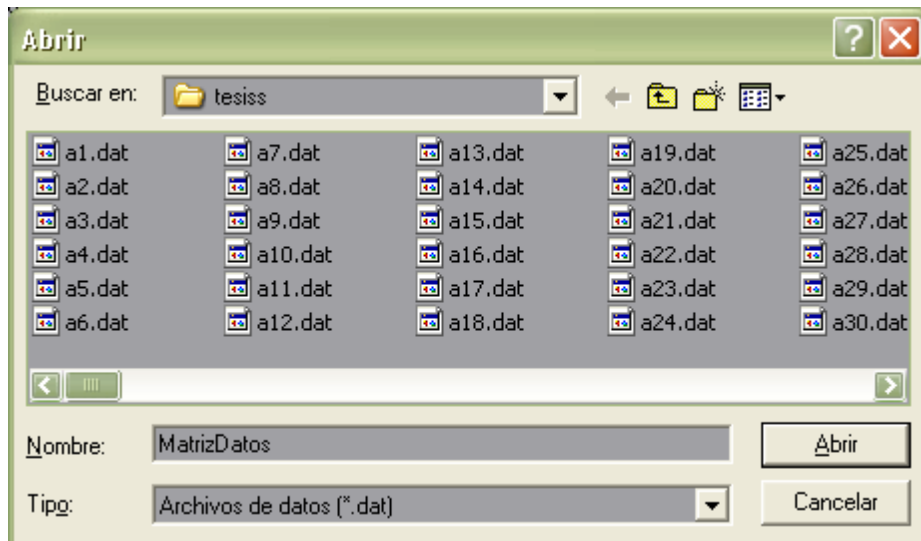


Figura 5.13

En esta ventana podrá seleccionar la ubicación y el nombre del archivo donde desea guardar la información mencionada anteriormente. El archivo guardado es un archivo de texto que incluye una identificación de que fue guardado con este sistema, así como del número de filas y columnas de la matriz de datos.

5.4 Exportar

Para guardar la información procesada hasta el momento en un archivo de texto se debe escoger la opción “Exportar...” del menú “Archivo” (vea la figura 5.14).

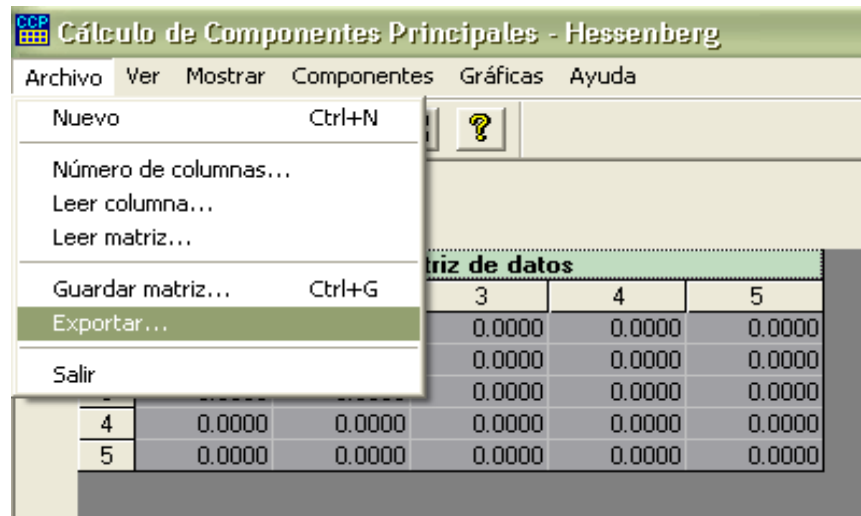


Figura 5.14

Al hacerlo aparecerá una ventana donde podrá seleccionar la ubicación y el nombre del archivo de texto que guardará la información (vea la figura 5.15).

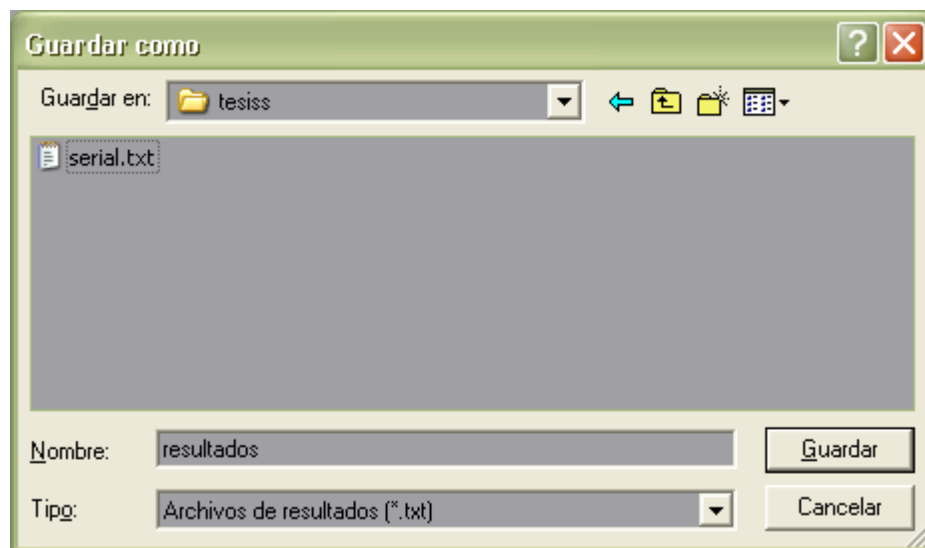


Figura 5.15

El archivo resultante puede incluso ser abierto con aplicaciones de otros paquetes, como Excel, para poder utilizar datos de otros archivos para mismas tareas.

5.5 Mostrar matrices

Las opciones de “Mostrar matrices”, “Datos estándar” y “Datos centrados” se realizan a través del menú “Mostrar” (ver figura 5.16).

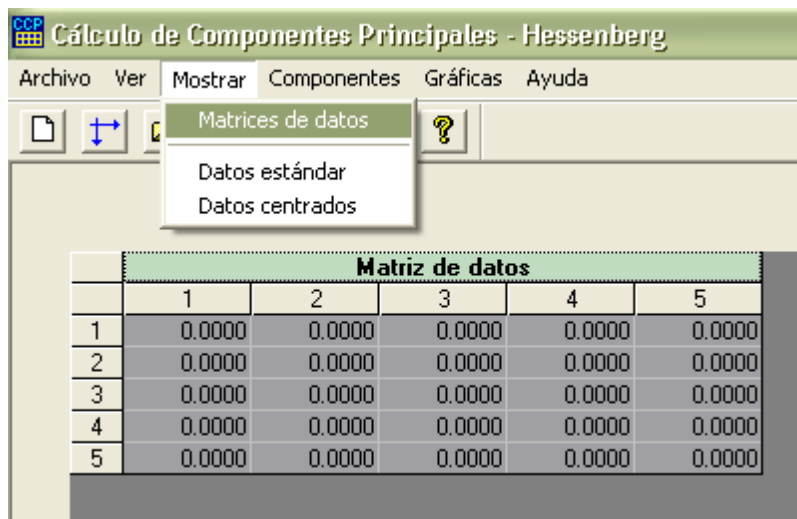


Figura 5.16

Al seleccionar estas opciones, se mostrará el resultado en las matrices designadas para la media, varianza, covarianza, correlación y matriz de resultados como se ve en la figura siguiente (figura 5.17):

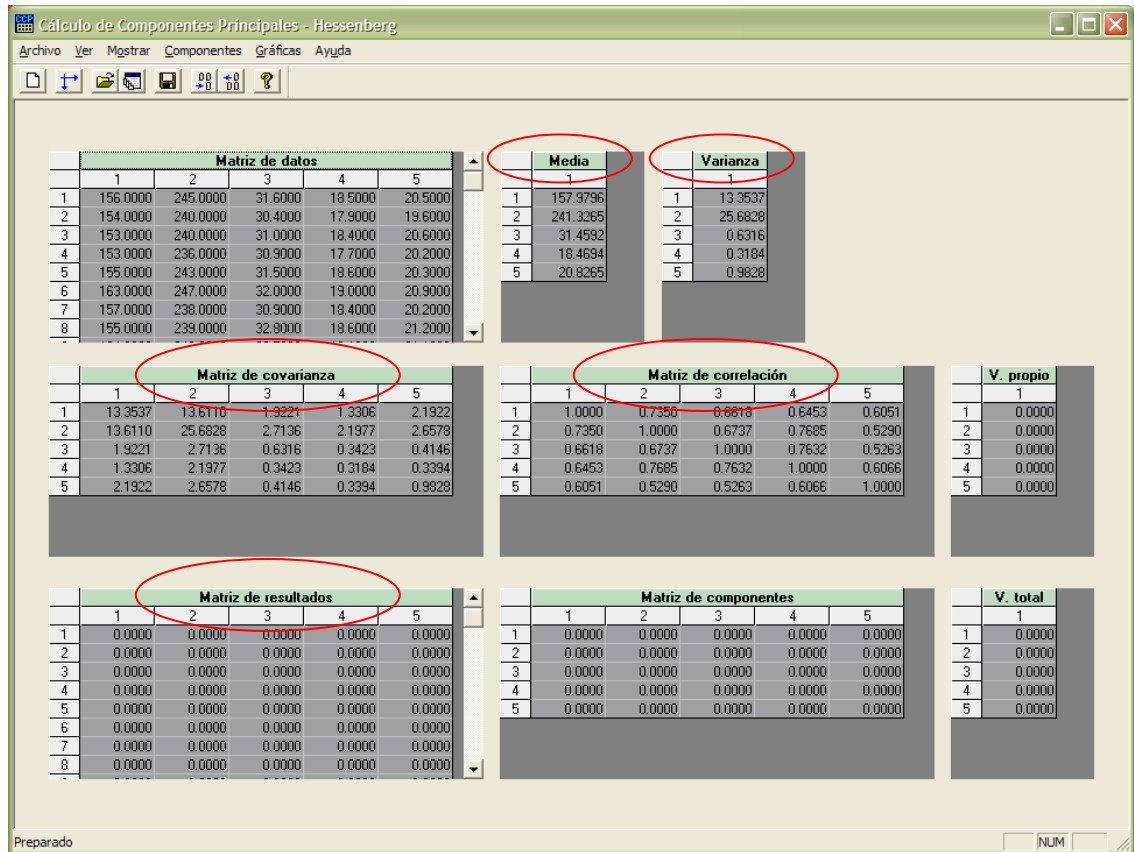


Figura 5.17

La matriz de resultados está designada a los datos estándar y datos centrados, por lo que para poder visualizar los resultados, se deberá de seleccionar uno sólo a la vez.

5.6 Cálculo de componentes principales

Las tareas de cálculo de componentes principales se dividen en seis partes, muy similares y ubicadas en el menú “Componentes”. El resultado de las mismas se ve reflejado en la matriz de componentes, el vector de valores propios y el vector de varianza total. Para ejecutar alguna de ellas selecciónela del menú como se muestra en la figura 5.18 (por ejemplo Correlación/Datos originales).

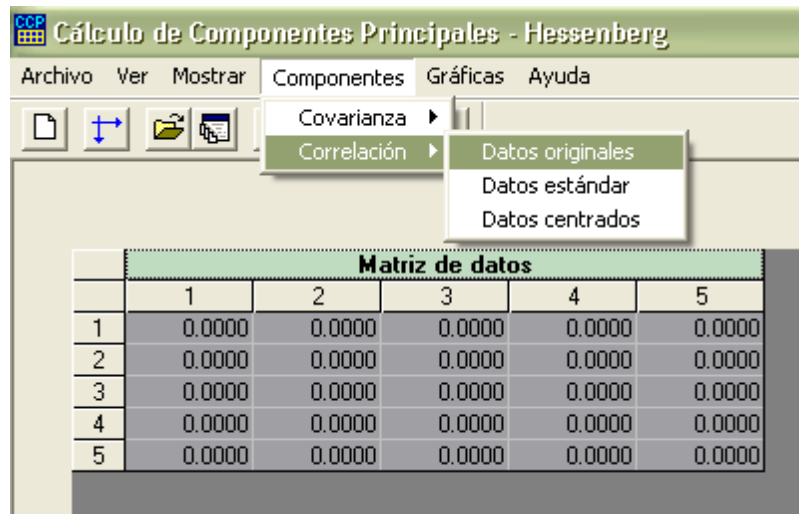


Figura 5.18

En la siguiente figura (5.19) se muestran los lugares donde se reflejan los resultados. Aquí, al igual que en las matrices de datos, los resultados comparten las mismas matrices y vectores, por lo que no se podrán visualizar simultáneamente.

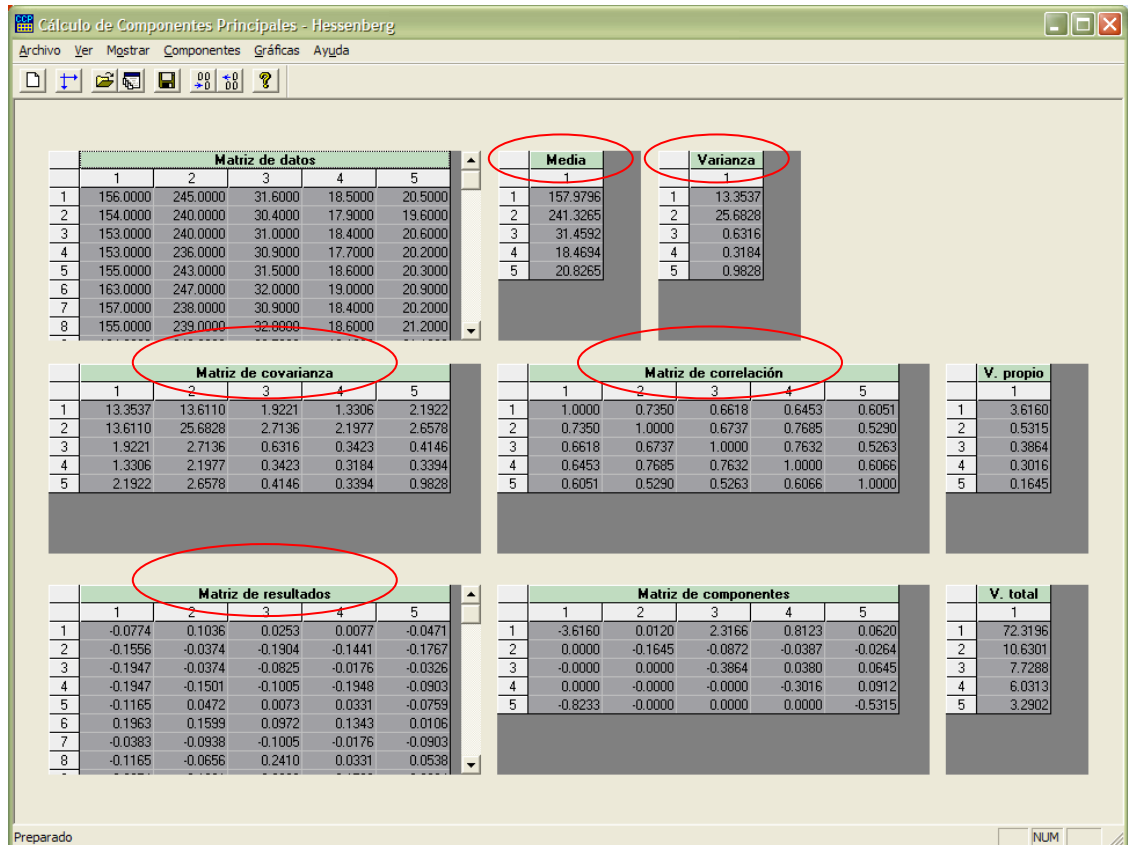


Figura 5.19

5.7 Gráficas

Las gráficas disponibles en este sistema son:

- Datos originales.
- Valores propios.

Ambos tipos de gráficas se seleccionan del menú "Gráficas" como se ve en la figura 5.20 (por ejemplo Gráficas/Datos originales).

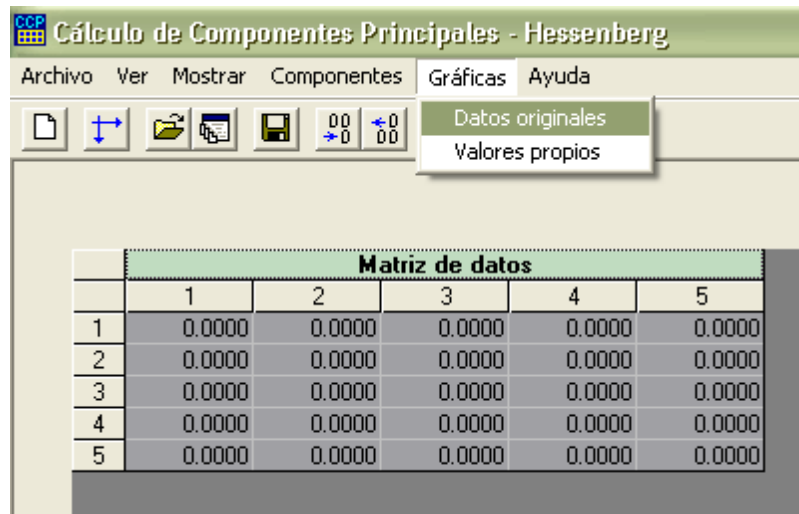


Figura 5.20

Ambas gráficas se muestran en una ventana separada.

5.7.1 Datos originales

Esta gráfica muestra los valores por columna de la matriz de datos. Se cuenta con una lista que permite seleccionar la columna a visualizar y presenta los valores Media y Varianza (ver figura 5.21).

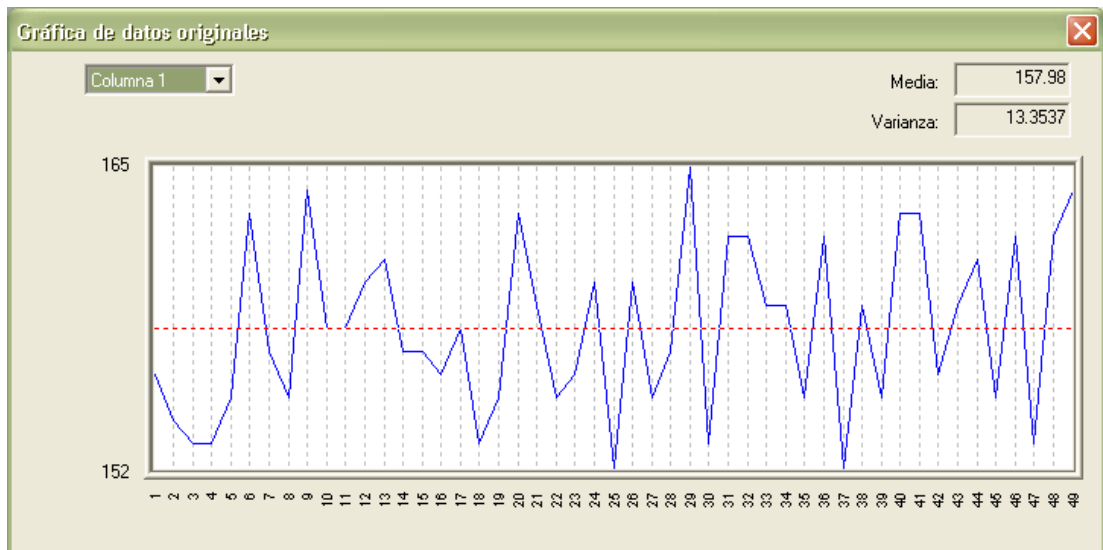


Figura 5.21

5.7.2 Valores Propios

Esta gráfica muestra (en forma similar a la anterior) los valores del vector de valores propios (ver figura 5.22). En el eje X se muestra el número de valores propios y el eje Y representa el valor correspondiente a cada uno de ellos.

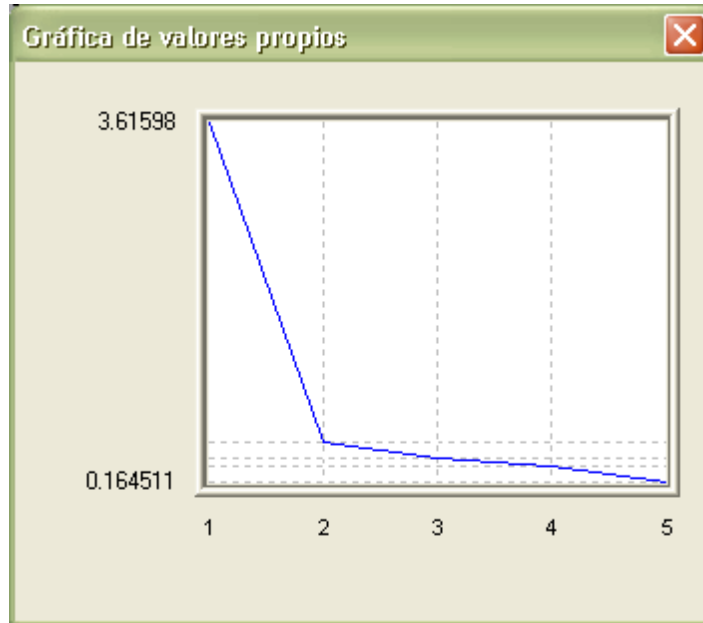


Figura 5.22

Puesto que algunos valores están muy “cercaños”, solo se mostrarán el valor máximo y el mínimo en el eje Y.

5.8 Barra de herramientas

Como se mencionó anteriormente, esta aplicación incluye algunos botones que ayudan a acelerar la selección de opciones de los menús. Estos se encuentran en la barra de herramientas y a continuación se indican las funciones de cada botón (vea la figura 5.23) en el orden en que aparecen.

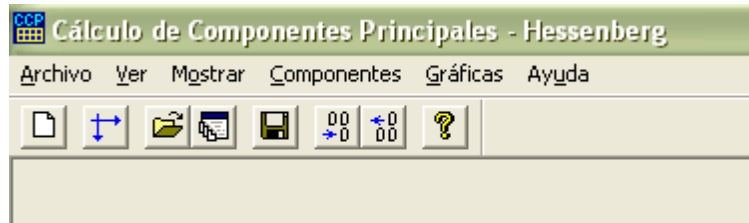










Figura 5.23

-  “Nuevo”. Similar a la opción Nuevo del menú Archivo.
-  “Dimensión de la matriz”. Similar a la opción “Número de columnas...” del menú archivo.
-  “Lee columna de datos”. Similar a la opción “Leer columna...” del menú archivo.
-  “Lee matriz de datos”. Similar a la opción “Leer matriz...” del menú archivo.
-  “Guardar matriz de datos”. Similar a la opción “Guardar matriz...” del menú archivo.
-  “Menos decimales”. Disminuye el número de decimales que se mostrarán en las matrices y vectores.
-  “Más decimales”. Incrementa el número de decimales que se mostrarán en las matrices y vectores.
-  “Acerca de...”. Muestra información sobre la versión del programa.

Note que cuando el número de decimales es cero y se hace clic en el botón “Menos decimales”, se mostrarán los decimales que Visual C++ juzgue más conveniente. Desde aquí se puede presionar el botón “Más decimales” para volver a mostrar la cantidad de decimales que se desee (0, 1, 2, 3,...).

Capítulo 6. Pruebas

En este capítulo, ya se presentan resultados de las pruebas realizadas al sistema que comprueban su funcionamiento. Se hace la descripción de dos ejemplos con datos que permitieron ver la generalidad por el sistema creado y se muestran los resultados obtenidos.

6.1 Ejemplo 1

Se utilizó el experimento realizado por Hermon Bumpus en 1898, en la Universidad de Brown, Rhode Island, el cual consistió en tomar las medidas del cuerpo de 49 gorriones [3]. Estas medidas dieron origen a 5 variables X1, X2, X3, X4 y X5, las cuales se describen a continuación:

- X1 = Longitud total.
- X2 = Longitud de ala.
- X3 = Longitud del pico y cabeza.
- X4 = Longitud del húmero.
- X5 = Longitud de la quilla del esternón.

Las medidas tomadas se encuentran en milímetros. Los datos leídos se presentan a continuación:

| | | | | |
|-----|-----|------|------|------|
| 156 | 245 | 31.6 | 18.5 | 20.5 |
| 154 | 240 | 30.4 | 17.9 | 19.6 |
| 153 | 240 | 31 | 18.4 | 20.6 |
| 153 | 236 | 30.9 | 17.7 | 20.2 |
| 155 | 243 | 31.5 | 18.6 | 20.3 |
| 163 | 247 | 32 | 19 | 20.9 |
| 157 | 238 | 30.9 | 18.4 | 20.2 |
| 155 | 239 | 32.8 | 18.6 | 21.2 |
| 164 | 248 | 32.7 | 19.1 | 21.1 |
| 158 | 238 | 31 | 18.8 | 22 |
| 158 | 240 | 31.3 | 18.6 | 22 |
| 160 | 244 | 31.1 | 18.6 | 20.5 |
| 161 | 246 | 32.3 | 19.3 | 21.8 |

| | | | | |
|-----|-----|------|------|------|
| 157 | 245 | 32 | 19.1 | 20 |
| 157 | 235 | 31.5 | 18.1 | 19.8 |
| 156 | 237 | 30.9 | 18 | 20.3 |
| 158 | 244 | 31.4 | 18.5 | 21.6 |
| 153 | 238 | 30.5 | 18.2 | 20.9 |
| 155 | 236 | 30.3 | 18.5 | 20.1 |
| 163 | 246 | 32.5 | 18.6 | 21.9 |
| 159 | 236 | 31.5 | 18 | 21.5 |
| 155 | 240 | 31.4 | 18 | 20.7 |
| 156 | 240 | 31.5 | 18.2 | 20.6 |
| 160 | 242 | 32.6 | 18.8 | 21.7 |
| 152 | 232 | 30.3 | 17.2 | 19.8 |
| 160 | 250 | 31.7 | 18.8 | 22.5 |
| 155 | 237 | 31 | 18.5 | 20 |
| 157 | 245 | 32.2 | 19.5 | 21.4 |
| 165 | 245 | 33.1 | 19.8 | 22.7 |
| 153 | 231 | 30.1 | 17.3 | 19.8 |
| 162 | 239 | 30.3 | 18 | 23.1 |
| 162 | 243 | 31.6 | 18.8 | 21.3 |
| 159 | 245 | 31.8 | 18.5 | 21.7 |
| 159 | 247 | 30.9 | 18.1 | 19 |
| 155 | 243 | 30.9 | 18.5 | 21.3 |
| 162 | 252 | 31.9 | 19.1 | 22.2 |
| 152 | 230 | 30.4 | 17.3 | 18.6 |
| 159 | 242 | 30.8 | 18.2 | 20.5 |
| 155 | 238 | 31.2 | 17.9 | 19.3 |
| 163 | 249 | 33.4 | 19.5 | 22.8 |
| 163 | 242 | 31 | 18.1 | 20.7 |
| 156 | 237 | 31.7 | 18.2 | 20.3 |
| 159 | 238 | 31.5 | 18.4 | 20.3 |
| 161 | 245 | 32.1 | 19.1 | 20.8 |
| 155 | 235 | 30.7 | 17.7 | 19.6 |
| 162 | 247 | 31.9 | 19.1 | 20.4 |
| 153 | 237 | 30.6 | 18.6 | 20.4 |
| 162 | 245 | 32.5 | 18.5 | 21.1 |
| 164 | 248 | 32.3 | 18.8 | 20.9 |

Posteriormente se le aplicaron las tareas que realiza el sistema obteniéndose los resultados siguientes (mostrados con 3 decimales):

Matriz de correlación:

| | | | | |
|-------|-------|-------|-------|-------|
| 1.000 | 0.735 | 0.662 | 0.645 | 0.605 |
| 0.735 | 1.000 | 0.674 | 0.769 | 0.529 |
| 0.662 | 0.674 | 1.000 | 0.763 | 0.526 |

| | | | | |
|-------|-------|-------|-------|-------|
| 0.645 | 0.769 | 0.763 | 1.000 | 0.607 |
| 0.605 | 0.529 | 0.526 | 0.607 | 1.000 |

Valores propios:

| |
|-------|
| 3.616 |
| 0.532 |
| 0.386 |
| 0.302 |
| 0.165 |

Los datos anteriores concuerdan con los obtenidos en el experimento mencionado anteriormente.

A continuación mostramos el archivo guardado con la opción "Exportar" después de ejecutar la opción "Correlación/Datos originales" y una imagen del programa.

```

Matriz de datos
156 245 31.6 18.5 20.5
154 240 30.4 17.9 19.6
153 240 31 18.4 20.6
153 236 30.9 17.7 20.2
155 243 31.5 18.6 20.3
163 247 32 19 20.9
157 238 30.9 18.4 20.2
155 239 32.8 18.6 21.2
164 248 32.7 19.1 21.1
158 238 31 18.8 22
158 240 31.3 18.6 22
160 244 31.1 18.6 20.5
161 246 32.3 19.3 21.8
157 245 32 19.1 20
157 235 31.5 18.1 19.8
156 237 30.9 18 20.3
158 244 31.4 18.5 21.6
153 238 30.5 18.2 20.9
155 236 30.3 18.5 20.1
163 246 32.5 18.6 21.9
159 236 31.5 18 21.5
155 240 31.4 18 20.7
156 240 31.5 18.2 20.6
160 242 32.6 18.8 21.7
152 232 30.3 17.2 19.8
160 250 31.7 18.8 22.5
155 237 31 18.5 20
157 245 32.2 19.5 21.4
165 245 33.1 19.8 22.7
153 231 30.1 17.3 19.8
162 239 30.3 18 23.1

```

| | | | | |
|-----|-----|------|------|------|
| 162 | 243 | 31.6 | 18.8 | 21.3 |
| 159 | 245 | 31.8 | 18.5 | 21.7 |
| 159 | 247 | 30.9 | 18.1 | 19 |
| 155 | 243 | 30.9 | 18.5 | 21.3 |
| 162 | 252 | 31.9 | 19.1 | 22.2 |
| 152 | 230 | 30.4 | 17.3 | 18.6 |
| 159 | 242 | 30.8 | 18.2 | 20.5 |
| 155 | 238 | 31.2 | 17.9 | 19.3 |
| 163 | 249 | 33.4 | 19.5 | 22.8 |
| 163 | 242 | 31 | 18.1 | 20.7 |
| 156 | 237 | 31.7 | 18.2 | 20.3 |
| 159 | 238 | 31.5 | 18.4 | 20.3 |
| 161 | 245 | 32.1 | 19.1 | 20.8 |
| 155 | 235 | 30.7 | 17.7 | 19.6 |
| 162 | 247 | 31.9 | 19.1 | 20.4 |
| 153 | 237 | 30.6 | 18.6 | 20.4 |
| 162 | 245 | 32.5 | 18.5 | 21.1 |
| 164 | 248 | 32.3 | 18.8 | 20.9 |

Vector de medias

157.98
241.327
31.4592
18.4694
20.8265

Vector de varianzas

13.3537
25.6828
0.631633
0.318418
0.982823

Matriz de covarianza

| | | | | |
|---------|---------|----------|----------|----------|
| 13.3537 | 13.611 | 1.92207 | 1.33061 | 2.19222 |
| 13.611 | 25.6828 | 2.71361 | 2.1977 | 2.65782 |
| 1.92207 | 2.71361 | 0.631633 | 0.342266 | 0.414647 |
| 1.33061 | 2.1977 | 0.342266 | 0.318418 | 0.339371 |
| 2.19222 | 2.65782 | 0.414647 | 0.339371 | 0.982823 |

Matriz de correlación

| | | | | |
|----------|----------|----------|----------|----------|
| 1 | 0.734964 | 0.661812 | 0.645284 | 0.605125 |
| 0.734964 | 1 | 0.673741 | 0.768509 | 0.529014 |
| 0.661812 | 0.673741 | 1 | 0.76319 | 0.52627 |
| 0.645284 | 0.768509 | 0.76319 | 1 | 0.606649 |
| 0.605125 | 0.529014 | 0.52627 | 0.606649 | 1 |

Matriz de resultados

| | | | | |
|------------|------------|------------|------------|------------|
| -0.0773884 | 0.103552 | 0.0253118 | 0.00774994 | -0.0470531 |
| -0.155575 | -0.0373936 | -0.190389 | -0.144149 | -0.176743 |
| -0.194668 | -0.0373936 | -0.0825384 | -0.0175665 | -0.0326431 |
| -0.194668 | -0.15015 | -0.100513 | -0.194782 | -0.0902831 |
| -0.116482 | 0.0471735 | 0.00733675 | 0.0330664 | -0.0758731 |
| 0.196263 | 0.15993 | 0.0972119 | 0.134332 | 0.0105869 |

| | | | | |
|-------------|------------|------------|------------|-------------|
| -0.0382953 | -0.0937718 | -0.100513 | -0.0175665 | -0.0902831 |
| -0.116482 | -0.0655827 | 0.241012 | 0.0330664 | 0.053817 |
| 0.235357 | 0.188119 | 0.223037 | 0.159649 | 0.039407 |
| 0.000797819 | -0.0937718 | -0.0825384 | 0.0836993 | 0.169097 |
| 0.000797819 | -0.0373936 | -0.0286133 | 0.0330664 | 0.169097 |
| 0.0789841 | 0.0753626 | -0.0645634 | 0.0330664 | -0.0470531 |
| 0.118077 | 0.131741 | 0.151137 | 0.210282 | 0.140277 |
| -0.0382953 | 0.103552 | 0.0972119 | 0.159649 | -0.119103 |
| -0.0382953 | -0.178339 | 0.00733675 | -0.0935159 | -0.147923 |
| -0.0773884 | -0.121961 | -0.100513 | -0.118832 | -0.0758731 |
| 0.000797819 | 0.0753626 | -0.0106383 | 0.00774994 | 0.111457 |
| -0.194668 | -0.0937718 | -0.172414 | -0.0681994 | 0.0105869 |
| -0.116482 | -0.15015 | -0.208364 | 0.00774994 | -0.104693 |
| 0.196263 | 0.131741 | 0.187087 | 0.0330664 | 0.154687 |
| 0.039891 | -0.15015 | 0.00733675 | -0.118832 | 0.097047 |
| -0.116482 | -0.0373936 | -0.0106383 | -0.118832 | -0.0182331 |
| -0.0773884 | -0.0373936 | 0.00733675 | -0.0681994 | -0.0326431 |
| 0.0789841 | 0.0189845 | 0.205062 | 0.0836993 | 0.125867 |
| -0.233761 | -0.262906 | -0.208364 | -0.321364 | -0.147923 |
| 0.0789841 | 0.244497 | 0.0432868 | 0.0836993 | 0.241147 |
| -0.116482 | -0.121961 | -0.0825384 | 0.00774994 | -0.119103 |
| -0.0382953 | 0.103552 | 0.133162 | 0.260914 | 0.082637 |
| 0.27445 | 0.103552 | 0.294937 | 0.336864 | 0.269967 |
| -0.194668 | -0.291095 | -0.244314 | -0.296048 | -0.147923 |
| 0.15717 | -0.0655827 | -0.208364 | -0.118832 | 0.327607 |
| 0.15717 | 0.0471735 | 0.0253118 | 0.0836993 | 0.068227 |
| 0.039891 | 0.103552 | 0.0612618 | 0.00774994 | 0.125867 |
| 0.039891 | 0.15993 | -0.100513 | -0.0935159 | -0.263203 |
| -0.116482 | 0.0471735 | -0.100513 | 0.00774994 | 0.068227 |
| 0.15717 | 0.300875 | 0.0792369 | 0.159649 | 0.197917 |
| -0.233761 | -0.319284 | -0.190389 | -0.296048 | -0.320843 |
| 0.039891 | 0.0189845 | -0.118488 | -0.0681994 | -0.0470531 |
| -0.116482 | -0.0937718 | -0.0465883 | -0.144149 | -0.219973 |
| 0.196263 | 0.216308 | 0.348862 | 0.260914 | 0.284377 |
| 0.196263 | 0.0189845 | -0.0825384 | -0.0935159 | -0.0182331 |
| -0.0773884 | -0.121961 | 0.0432868 | -0.0681994 | -0.0758731 |
| 0.039891 | -0.0937718 | 0.00733675 | -0.0175665 | -0.0758731 |
| 0.118077 | 0.103552 | 0.115187 | 0.159649 | -0.00382306 |
| -0.116482 | -0.178339 | -0.136464 | -0.194782 | -0.176743 |
| 0.15717 | 0.15993 | 0.0792369 | 0.159649 | -0.0614631 |
| -0.194668 | -0.121961 | -0.154439 | 0.0330664 | -0.0614631 |
| 0.15717 | 0.103552 | 0.187087 | 0.00774994 | 0.039407 |
| 0.235357 | 0.188119 | 0.151137 | 0.0836993 | 0.0105869 |

Matriz de componentes

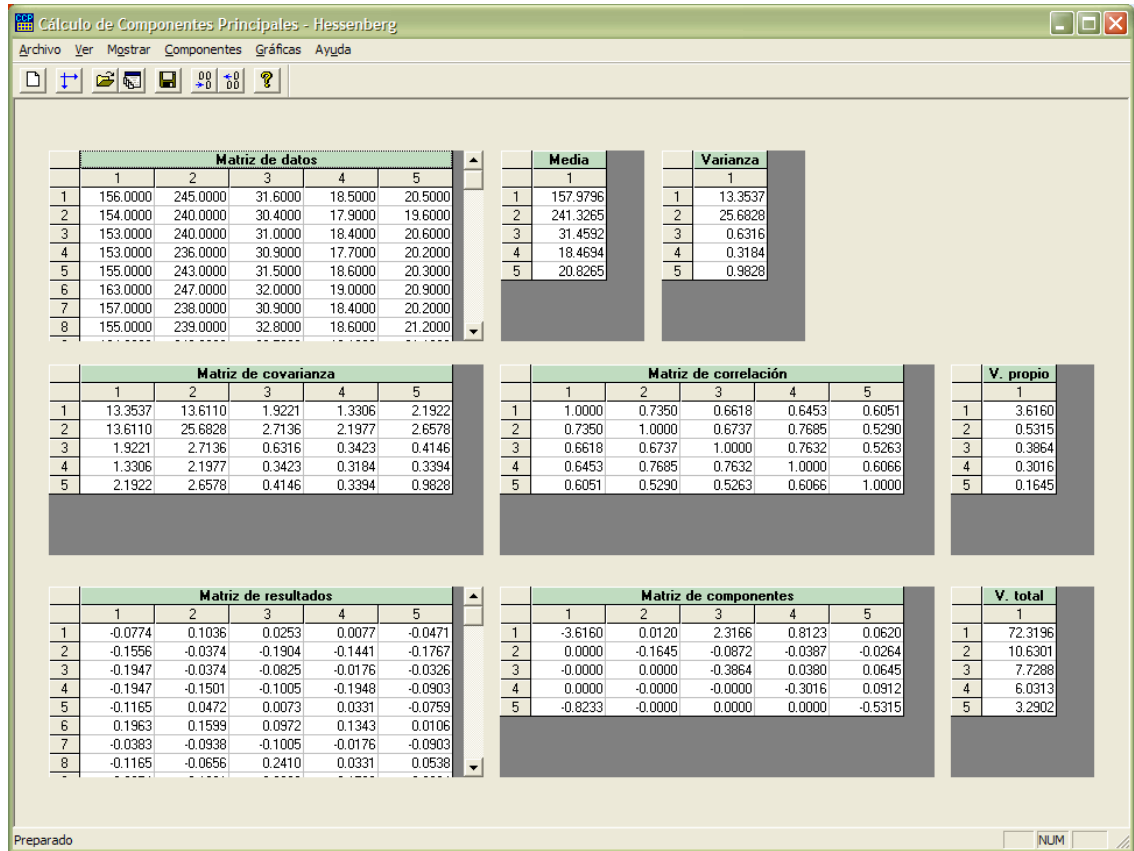
| | | | | |
|---------------|---------------|---------------|--------------|------------|
| -3.61598 | 0.0120085 | 2.31662 | 0.812261 | 0.0619532 |
| 2.09017e-026 | -0.164511 | -0.0872296 | -0.0387319 | -0.0263508 |
| -2.78073e-029 | 4.28848e-005 | -0.386441 | 0.0380201 | 0.0645342 |
| 2.98892e-026 | -2.59926e-016 | -6.35128e-026 | -0.301566 | 0.0911533 |
| -0.823339 | -4.41027e-023 | 6.84909e-033 | 1.63123e-010 | -0.531504 |

Vector de valores propios

3.61598
0.531504
0.386441
0.301566
0.164511

Vector de varianza total

72.3196
 10.6301
 7.72883
 6.03131
 3.29021



6.2 Ejemplo 2

En este ejemplo se tomaron porcentajes de la fuerza laboral en nueve sectores diferentes de la industria en 26 naciones europeas. Los datos son los siguientes (fuente: Euromonitor. 1979. pp 76-7):

| | | | | | | | | |
|------|-----|------|-----|-----|------|-----|------|-----|
| 3.3 | 0.9 | 27.6 | 0.9 | 8.2 | 19.1 | 6.2 | 26.6 | 7.2 |
| 9.2 | 0.1 | 21.8 | 0.6 | 8.3 | 14.6 | 6.5 | 32.2 | 7.1 |
| 10.8 | 0.8 | 27.5 | 0.9 | 8.9 | 16.8 | 6 | 22.6 | 5.7 |
| 6.7 | 1.3 | 35.8 | 0.9 | 7.3 | 14.4 | 5 | 22.3 | 6.1 |
| 23.2 | 1 | 20.7 | 1.3 | 7.5 | 16.8 | 2.8 | 20.8 | 6.1 |
| 15.9 | 0.6 | 27.6 | 0.5 | 10 | 18.1 | 1.6 | 20.1 | 5.7 |
| 7.7 | 3.1 | 30.8 | 0.8 | 9.2 | 18.5 | 4.6 | 19.2 | 6.2 |
| 6.3 | 0.1 | 22.5 | 1 | 9.9 | 18 | 6.8 | 28.5 | 6.8 |

| | | | | | | | | |
|------|-----|------|-----|------|------|------|------|-----|
| 2.7 | 1.4 | 30.2 | 1.4 | 6.9 | 16.9 | 5.7 | 28.3 | 6.4 |
| 12.7 | 1.1 | 30.2 | 1.4 | 9 | 16.8 | 4.9 | 16.8 | 7 |
| 13 | 0.4 | 25.9 | 1.3 | 7.4 | 14.7 | 5.5 | 24.3 | 7.6 |
| 41.4 | 0.6 | 17.6 | 0.6 | 8.1 | 11.5 | 2.4 | 11 | 6.7 |
| 9 | 0.5 | 22.4 | 0.8 | 8.6 | 16.9 | 4.7 | 27.6 | 9.4 |
| 27.8 | 0.3 | 24.5 | 0.6 | 8.4 | 13.3 | 2.7 | 16.7 | 5.7 |
| 22.9 | 0.8 | 28.5 | 0.7 | 11.5 | 9.7 | 8.5 | 11.8 | 5.5 |
| 6.1 | 0.4 | 25.9 | 0.8 | 7.2 | 14.4 | 6 | 32.4 | 6.8 |
| 7.7 | 0.2 | 37.8 | 0.8 | 9.5 | 17.5 | 5.3 | 15.4 | 5.7 |
| 66.8 | 0.7 | 7.9 | 0.1 | 2.8 | 5.2 | 1.1 | 11.9 | 3.2 |
| 23.6 | 1.9 | 32.3 | 0.6 | 7.9 | 8 | 0.7 | 18.2 | 6.7 |
| 16.5 | 2.9 | 35.5 | 1.2 | 8.7 | 9.2 | 0.9 | 17.9 | 7 |
| 4.2 | 2.9 | 41.2 | 1.3 | 7.6 | 11.2 | 1.2 | 22.1 | 8.4 |
| 21.7 | 3.1 | 29.6 | 1.9 | 8.2 | 9.4 | 0.9 | 17.2 | 8 |
| 31.1 | 2.5 | 25.7 | 0.9 | 8.4 | 7.5 | 0.9 | 16.1 | 6.9 |
| 34.7 | 2.1 | 30.1 | 0.6 | 8.7 | 5.9 | 1.3 | 11.7 | 5 |
| 23.7 | 1.4 | 25.8 | 0.6 | 9.2 | 6.1 | 0.5 | 23.6 | 9.3 |
| 48.7 | 1.5 | 16.8 | 1.1 | 4.9 | 6.4 | 11.3 | 5.3 | 4 |

Al utilizar el sistema de cálculo de componentes principales se obtuvieron los siguientes resultados:

Matriz de correlación:

| | | | | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1.000 | 0.036 | -0.671 | -0.400 | -0.538 | -0.737 | -0.220 | -0.747 | -0.565 |
| 0.036 | 1.000 | 0.445 | 0.405 | -0.026 | -0.397 | -0.443 | -0.281 | 0.157 |
| -0.671 | 0.445 | 1.000 | 0.385 | 0.494 | 0.204 | -0.156 | 0.154 | 0.351 |
| -0.400 | 0.405 | 0.385 | 1.000 | 0.060 | 0.202 | 0.110 | 0.132 | 0.375 |
| -0.538 | -0.026 | 0.494 | 0.060 | 1.000 | 0.356 | 0.016 | 0.158 | 0.388 |
| -0.737 | -0.397 | 0.204 | 0.202 | 0.356 | 1.000 | 0.366 | 0.572 | 0.188 |
| -0.220 | -0.443 | -0.156 | 0.110 | 0.016 | 0.366 | 1.000 | 0.108 | -0.246 |
| -0.747 | -0.281 | 0.154 | 0.132 | 0.158 | 0.572 | 0.108 | 1.000 | 0.568 |
| -0.565 | 0.157 | 0.351 | 0.375 | 0.388 | 0.188 | -0.246 | 0.568 | 1.000 |

Valores propios:

| |
|-------|
| 3.487 |
| 2.130 |
| 1.099 |
| 0.995 |
| 0.543 |
| 0.383 |
| 0.226 |
| 0.137 |
| 0.000 |

Los resultados también concuerdan con los obtenidos y a continuación mostramos el archivo exportado con los resultados de la opción “Correlación/Datos originales” y unas imágenes del programa.

Matriz de datos

| | | | | | | | | |
|------|-----|------|-----|------|------|------|------|-----|
| 3.3 | 0.9 | 27.6 | 0.9 | 8.2 | 19.1 | 6.2 | 26.6 | 7.2 |
| 9.2 | 0.1 | 21.8 | 0.6 | 8.3 | 14.6 | 6.5 | 32.2 | 7.1 |
| 10.8 | 0.8 | 27.5 | 0.9 | 8.9 | 16.8 | 6 | 22.6 | 5.7 |
| 6.7 | 1.3 | 35.8 | 0.9 | 7.3 | 14.4 | 5 | 22.3 | 6.1 |
| 23.2 | 1 | 20.7 | 1.3 | 7.5 | 16.8 | 2.8 | 20.8 | 6.1 |
| 15.9 | 0.6 | 27.6 | 0.5 | 10 | 18.1 | 1.6 | 20.1 | 5.7 |
| 7.7 | 3.1 | 30.8 | 0.8 | 9.2 | 18.5 | 4.6 | 19.2 | 6.2 |
| 6.3 | 0.1 | 22.5 | 1 | 9.9 | 18 | 6.8 | 28.5 | 6.8 |
| 2.7 | 1.4 | 30.2 | 1.4 | 6.9 | 16.9 | 5.7 | 28.3 | 6.4 |
| 12.7 | 1.1 | 30.2 | 1.4 | 9 | 16.8 | 4.9 | 16.8 | 7 |
| 13 | 0.4 | 25.9 | 1.3 | 7.4 | 14.7 | 5.5 | 24.3 | 7.6 |
| 41.4 | 0.6 | 17.6 | 0.6 | 8.1 | 11.5 | 2.4 | 11 | 6.7 |
| 9 | 0.5 | 22.4 | 0.8 | 8.6 | 16.9 | 4.7 | 27.6 | 9.4 |
| 27.8 | 0.3 | 24.5 | 0.6 | 8.4 | 13.3 | 2.7 | 16.7 | 5.7 |
| 22.9 | 0.8 | 28.5 | 0.7 | 11.5 | 9.7 | 8.5 | 11.8 | 5.5 |
| 6.1 | 0.4 | 25.9 | 0.8 | 7.2 | 14.4 | 6 | 32.4 | 6.8 |
| 7.7 | 0.2 | 37.8 | 0.8 | 9.5 | 17.5 | 5.3 | 15.4 | 5.7 |
| 66.8 | 0.7 | 7.9 | 0.1 | 2.8 | 5.2 | 1.1 | 11.9 | 3.2 |
| 23.6 | 1.9 | 32.3 | 0.6 | 7.9 | 8 | 0.7 | 18.2 | 6.7 |
| 16.5 | 2.9 | 35.5 | 1.2 | 8.7 | 9.2 | 0.9 | 17.9 | 7 |
| 4.2 | 2.9 | 41.2 | 1.3 | 7.6 | 11.2 | 1.2 | 22.1 | 8.4 |
| 21.7 | 3.1 | 29.6 | 1.9 | 8.2 | 9.4 | 0.9 | 17.2 | 8 |
| 31.1 | 2.5 | 25.7 | 0.9 | 8.4 | 7.5 | 0.9 | 16.1 | 6.9 |
| 34.7 | 2.1 | 30.1 | 0.6 | 8.7 | 5.9 | 1.3 | 11.7 | 5 |
| 23.7 | 1.4 | 25.8 | 0.6 | 9.2 | 6.1 | 0.5 | 23.6 | 9.3 |
| 48.7 | 1.5 | 16.8 | 1.1 | 4.9 | 6.4 | 11.3 | 5.3 | 4 |

Vector de medias

19.1308
 1.25385
 27.0077
 0.907692
 8.16538
 12.9577
 4
 20.0231
 6.54615

Vector de varianzas

241.696
 0.940985
 49.1087
 0.141538
 2.70795
 20.9329
 7.8768
 46.6426
 1.93618

Matriz de covarianza

| | | | | | | | | |
|----------|------------|----------|-----------|------------|----------|---------|----------|----------|
| 241.696 | 0.539877 | -73.1138 | -2.33985 | -13.7721 | -52.421 | -9.592 | -79.2911 | -12.2207 |
| 0.539877 | 0.940985 | 3.02637 | 0.147969 | -0.0408615 | -1.76003 | -1.2052 | -1.86169 | 0.211415 |
| -73.1138 | 3.02637 | 49.1087 | 1.01594 | 5.70228 | 6.53514 | -3.0648 | 7.37862 | 3.41963 |
| -2.33985 | 0.147969 | 1.01594 | 0.141538 | 0.0370769 | 0.347538 | 0.116 | 0.340215 | 0.196431 |
| -13.7721 | -0.0408615 | 5.70228 | 0.0370769 | 2.70795 | 2.68048 | 0.0752 | 1.77843 | 0.887662 |
| -52.421 | -1.76003 | 6.53514 | 0.347538 | 2.68048 | 20.9329 | 4.694 | 17.8786 | 1.19403 |
| -9.592 | -1.2052 | -3.0648 | 0.116 | 0.0752 | 4.694 | 7.8768 | 2.0632 | -0.9604 |
| -79.2911 | -1.86169 | 7.37862 | 0.340215 | 1.77843 | 17.8786 | 2.0632 | 46.6426 | 5.39649 |
| -12.2207 | 0.211415 | 3.41963 | 0.196431 | 0.887662 | 1.19403 | -0.9604 | 5.39649 | 1.93618 |

Matriz de correlación

| | | | | | | | | |
|-----------|------------|-----------|-----------|------------|-----------|-----------|-----------|-----------|
| 1 | 0.0357988 | -0.671098 | -0.400051 | -0.538325 | -0.736981 | -0.219836 | -0.74679 | -0.56492 |
| 0.0357988 | 1 | 0.445196 | 0.405455 | -0.0255978 | -0.396565 | -0.442683 | -0.281012 | 0.156629 |
| -0.671098 | 0.445196 | 1 | 0.385346 | 0.494479 | 0.203826 | -0.155829 | 0.154171 | 0.350692 |
| -0.400051 | 0.405455 | 0.385346 | 1 | 0.0598888 | 0.201907 | 0.109862 | 0.132411 | 0.375231 |
| -0.538325 | -0.0255978 | 0.494479 | 0.0598888 | 1 | 0.356022 | 0.0162825 | 0.158243 | 0.387662 |
| -0.736981 | -0.396565 | 0.203826 | 0.201907 | 0.356022 | 1 | 0.365555 | 0.572173 | 0.187554 |
| -0.219836 | -0.442683 | -0.155829 | 0.109862 | 0.0162825 | 0.365555 | 1 | 0.10764 | -0.245926 |
| -0.74679 | -0.281012 | 0.154171 | 0.132411 | 0.158243 | 0.572173 | 0.10764 | 1 | 0.567867 |
| -0.56492 | 0.156629 | 0.350692 | 0.375231 | 0.387662 | 0.187554 | -0.245926 | 0.567867 | 1 |

Matriz de resultados

| | | | | | | | | |
|------------|------------|------------|-------------|-------------|------------|------------|------------|------------|
| -0.199701 | -0.071538 | 0.0165761 | -0.00400989 | 0.00412536 | 0.263287 | 0.153731 | 0.188862 | 0.0921543 |
| -0.125274 | -0.233276 | -0.14574 | -0.160396 | 0.0160431 | 0.0703968 | 0.174694 | 0.349671 | 0.0780601 |
| -0.105091 | -0.0917552 | 0.0137775 | -0.00400989 | 0.0875493 | 0.164699 | 0.139755 | 0.0739985 | -0.119258 |
| -0.156811 | 0.00933104 | 0.246058 | -0.00400989 | -0.103134 | 0.0618239 | 0.0698777 | 0.0653838 | -0.0628817 |
| 0.0513323 | -0.0513207 | -0.176524 | 0.204505 | -0.0792986 | 0.164699 | -0.0838532 | 0.02231 | -0.0628817 |
| -0.0407554 | -0.13219 | 0.0165761 | -0.212524 | 0.218644 | 0.220423 | -0.167706 | 0.00220891 | -0.119258 |
| -0.144196 | 0.373242 | 0.10613 | -0.0561385 | 0.123302 | 0.237569 | 0.0419266 | -0.0236354 | -0.0487876 |
| -0.161857 | -0.233276 | -0.12615 | 0.0481187 | 0.208726 | 0.216136 | 0.195657 | 0.243422 | 0.0357775 |
| -0.20727 | 0.0295483 | 0.0893385 | 0.256633 | -0.150805 | 0.168985 | 0.118792 | 0.237679 | -0.0205992 |
| -0.0811226 | -0.0311035 | 0.0893385 | 0.256633 | 0.099467 | 0.164699 | 0.0628899 | -0.0925534 | 0.0639659 |
| -0.0773381 | -0.172624 | -0.0309994 | 0.204505 | -0.0912163 | 0.0746832 | 0.104817 | 0.122815 | 0.148531 |
| 0.280921 | -0.13219 | -0.263279 | -0.160396 | -0.00779235 | -0.0624833 | -0.111804 | -0.259105 | 0.0216834 |
| -0.127797 | -0.152407 | -0.128949 | -0.0561385 | 0.0517962 | 0.168985 | 0.0489144 | 0.217578 | 0.402226 |
| 0.10936 | -0.192841 | -0.0701792 | -0.160396 | 0.0279608 | 0.0146729 | -0.090841 | -0.095425 | -0.119258 |
| 0.0475479 | -0.0917552 | 0.0417631 | -0.108267 | 0.39741 | -0.139644 | 0.31445 | -0.236133 | -0.147447 |
| -0.16438 | -0.172624 | -0.0309994 | -0.0561385 | -0.115052 | 0.0618239 | 0.139755 | 0.355414 | 0.0357775 |
| -0.144196 | -0.213059 | 0.302029 | -0.0561385 | 0.159056 | 0.194704 | 0.090841 | -0.132756 | -0.119258 |
| 0.601336 | -0.111972 | -0.534739 | -0.421039 | -0.639431 | -0.33253 | -0.202645 | -0.233261 | -0.471613 |
| 0.0563782 | 0.130635 | 0.148108 | -0.160396 | -0.0316278 | -0.212509 | -0.230596 | -0.0523512 | 0.0216834 |
| -0.0331865 | 0.332807 | 0.237662 | 0.152376 | 0.0637139 | -0.161072 | -0.216621 | -0.060966 | 0.0639659 |
| -0.188348 | 0.332807 | 0.39718 | 0.204505 | -0.0673809 | -0.0753427 | -0.195657 | 0.0596406 | 0.261284 |
| 0.0324102 | 0.373242 | 0.0725472 | 0.517276 | 0.00412536 | -0.152499 | -0.216621 | -0.0810671 | 0.204908 |
| 0.150989 | 0.251938 | -0.0365965 | -0.00400989 | 0.0279608 | -0.233942 | -0.216621 | -0.112655 | 0.0498717 |
| 0.196402 | 0.171069 | 0.0865399 | -0.160396 | 0.0637139 | -0.302525 | -0.18867 | -0.239004 | -0.217918 |
| 0.0576397 | 0.0295483 | -0.0337979 | -0.160396 | 0.123302 | -0.293952 | -0.244572 | 0.102714 | 0.388132 |
| 0.373009 | 0.0497655 | -0.285668 | 0.100247 | -0.389159 | -0.281093 | 0.510107 | -0.422786 | -0.35886 |

Matriz de componentes

| | | | | | | | | |
|---------------|---------------|---------------|--------------|---------------|---------------|---------------|---------------|-------------|
| -3.48715 | -0.183564 | -0.320815 | 0.523333 | 0.646314 | -2.68409 | -2.056 | -0.305876 | 0.792239 |
| -6.63439e-007 | -2.13017 | 0.456465 | -0.252213 | 0.340348 | 0.268982 | -0.588376 | 0.36987 | -0.598962 |
| -1.29679e-013 | 1.97215e-031 | -1.09892 | 0.0225937 | -0.0341898 | 0.267141 | 0.0425396 | 0.382594 | -0.32542 |
| -7.18518e-016 | 1.44445e-034 | 0.000187197 | -0.994523 | -0.0355797 | -0.355061 | -0.0484129 | 0.409546 | -0.377123 |
| -0.720852 | -7.08039e-021 | -3.49296e-014 | 5.95748e-025 | -0.543218 | 0.0694629 | 0.142896 | 0.119183 | 0.170352 |
| -0.986864 | -0.0996413 | 3.12933e-020 | 9.83765e-031 | 6.83504e-009 | -4.5626e-005 | 0.102941 | 0.0690522 | 0.0295396 |
| -0.294375 | 0.381764 | 0.661623 | 2.8637e-018 | -7.97029e-016 | 6.69027e-023 | -0.383428 | -0.108267 | -0.00719865 |
| 0.047937 | 0.0457978 | 0.215474 | -0.819545 | 1.0219e-014 | -4.68684e-018 | -1.47341e-022 | -0.13679 | -0.0452886 |
| -0.756465 | 0.247967 | -0.692308 | -0.687925 | -0.0803768 | 2.68609e-021 | 2.71853e-024 | -3.60977e-021 | -0.225754 |

Vector de valores propios

- 3.48715
- 2.13017
- 1.09892
- 0.994523
- 0.543218
- 0.383428
- 0.225754
- 0.13679
- 4.5626e-005

Vector de varianza total

- 38.7461
- 23.6686
- 12.2102
- 11.0503
- 6.03575
- 4.26031
- 2.50838
- 1.51989
- 0.000506955

Cálculo de Componentes Principales - Hessenberg

Archivo Ver Mostrar Componentes Gráficas Ayuda

| Matriz de datos | | | | | |
|-----------------|---------|--------|---------|--------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | 3.3000 | 0.9000 | 27.6000 | 0.9000 | 8.2000 |
| 2 | 9.2000 | 0.1000 | 21.8000 | 0.6000 | 8.3000 |
| 3 | 10.8000 | 0.8000 | 27.5000 | 0.9000 | 8.9000 |
| 4 | 6.7000 | 1.3000 | 35.8000 | 0.9000 | 7.3000 |
| 5 | 23.2000 | 1.0000 | 20.7000 | 1.3000 | 7.5000 |
| 6 | 15.9000 | 0.6000 | 27.6000 | 0.5000 | 10.0000 |
| 7 | 7.7000 | 3.1000 | 30.8000 | 0.8000 | 9.2000 |

| Media | |
|-------|---------|
| | 1 |
| 1 | 19.1308 |
| 2 | 1.2538 |
| 3 | 27.0077 |
| 4 | 0.9077 |
| 5 | 8.1654 |
| 6 | 12.9577 |
| 7 | 4.0000 |
| 8 | 20.0231 |

| Varianza | |
|----------|----------|
| | 1 |
| 1 | 241.6958 |
| 2 | 0.9410 |
| 3 | 49.1087 |
| 4 | 0.1415 |
| 5 | 2.7080 |
| 6 | 20.9329 |
| 7 | 7.8768 |
| 8 | 46.6426 |

| Matriz de covarianza | | | | | |
|----------------------|----------|---------|----------|---------|----------|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | 241.6958 | 0.5399 | -73.1138 | -2.3398 | -13.7721 |
| 2 | 0.5399 | 0.9410 | 3.0264 | 0.1480 | -0.0409 |
| 3 | -73.1138 | 3.0264 | 49.1087 | 1.0159 | 5.7023 |
| 4 | -2.3398 | 0.1480 | 1.0159 | 0.1415 | 0.0371 |
| 5 | -13.7721 | -0.0409 | 5.7023 | 0.0371 | 2.7080 |
| 6 | -52.4210 | -1.7600 | 6.5351 | 0.3475 | 2.6805 |
| 7 | -9.5920 | -1.2052 | -3.0648 | 0.1160 | 0.0752 |

| Matriz de correlación | | | | | |
|-----------------------|---------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | 1.0000 | 0.0358 | -0.6711 | -0.4001 | -0.5383 |
| 2 | 0.0358 | 1.0000 | 0.4452 | 0.4055 | -0.0256 |
| 3 | -0.6711 | 0.4452 | 1.0000 | 0.3853 | 0.4945 |
| 4 | -0.4001 | 0.4055 | 0.3853 | 1.0000 | 0.0599 |
| 5 | -0.5383 | -0.0256 | 0.4945 | 0.0599 | 1.0000 |
| 6 | -0.7370 | -0.3966 | 0.2038 | 0.2019 | 0.3560 |
| 7 | -0.2198 | -0.4427 | -0.1558 | 0.1099 | 0.0163 |

| V. propio | |
|-----------|--------|
| | 1 |
| 1 | 3.4872 |
| 2 | 2.1302 |
| 3 | 1.0989 |
| 4 | 0.9945 |
| 5 | 0.5432 |
| 6 | 0.3834 |
| 7 | 0.2258 |
| 8 | 0.1368 |

| Matriz de resultados | | | | | |
|----------------------|---------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | -0.1997 | -0.0715 | 0.0166 | -0.0040 | 0.0041 |
| 2 | -0.1253 | -0.2333 | -0.1457 | -0.1604 | 0.0160 |
| 3 | -0.1051 | -0.0918 | 0.0138 | -0.0040 | 0.0875 |
| 4 | -0.1568 | 0.0093 | 0.2461 | -0.0040 | -0.1031 |
| 5 | 0.0513 | -0.0513 | -0.1765 | 0.2045 | -0.0793 |
| 6 | -0.0408 | -0.1322 | 0.0166 | -0.2125 | 0.2186 |
| 7 | -0.1442 | 0.3732 | 0.1061 | -0.0561 | 0.1233 |

| Matriz de componentes | | | | | |
|-----------------------|---------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | -3.4872 | -0.1836 | -0.3208 | 0.5233 | 0.6463 |
| 2 | -0.0000 | -2.1302 | 0.4565 | -0.2522 | 0.3403 |
| 3 | -0.0000 | 0.0000 | -1.0989 | 0.0226 | -0.0342 |
| 4 | -0.0000 | 0.0000 | 0.0002 | -0.9945 | -0.0356 |
| 5 | -0.7209 | -0.0000 | -0.0000 | 0.0000 | -0.5432 |
| 6 | -0.9869 | -0.0996 | 0.0000 | 0.0000 | 0.0000 |
| 7 | -0.2944 | 0.3818 | 0.6616 | 0.0000 | -0.0000 |

| V. total | |
|----------|---------|
| | 1 |
| 1 | 38.7461 |
| 2 | 23.6686 |
| 3 | 12.2102 |
| 4 | 11.0503 |
| 5 | 6.0358 |
| 6 | 4.2603 |
| 7 | 2.5084 |
| 8 | 1.5199 |

Preparado NUM

Cálculo de Componentes Principales - Hessenberg

Archivo Ver Mostrar Componentes Gráficas Ayuda

| Matriz de datos | | | | | |
|-----------------|---------|--------|---------|--------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | 3.3000 | 0.9000 | 27.6000 | 0.9000 | 8.2000 |
| 2 | 9.2000 | 0.1000 | 21.8000 | 0.6000 | 8.3000 |
| 3 | 10.8000 | 0.8000 | 27.5000 | 0.9000 | 8.9000 |
| 4 | 6.7000 | 1.3000 | 35.8000 | 0.9000 | 7.3000 |
| 5 | 23.2000 | 1.0000 | 20.7000 | 1.3000 | 7.5000 |
| 6 | 15.9000 | 0.6000 | 27.6000 | 0.5000 | 10.0000 |
| 7 | 7.7000 | 3.1000 | 30.8000 | 0.8000 | 9.2000 |

| Media | |
|-------|---------|
| | 1 |
| 2 | 1.2538 |
| 3 | 27.0077 |
| 4 | 0.9077 |
| 5 | 8.1654 |
| 6 | 12.9577 |
| 7 | 4.0000 |
| 8 | 20.0231 |
| 9 | 6.5462 |

| Varianza | |
|----------|---------|
| | 1 |
| 2 | 0.9410 |
| 3 | 49.1087 |
| 4 | 0.1415 |
| 5 | 2.7080 |
| 6 | 20.9329 |
| 7 | 7.8768 |
| 8 | 46.6426 |
| 9 | 1.9362 |

| Matriz de covarianza | | | | | |
|----------------------|----------|---------|---------|--------|--------|
| | 1 | 2 | 3 | 4 | 5 |
| 3 | -73.1138 | 3.0264 | 49.1087 | 1.0159 | 5.7023 |
| 4 | -2.3398 | 0.1480 | 1.0159 | 0.1415 | 0.0371 |
| 5 | -13.7721 | -0.0409 | 5.7023 | 0.0371 | 2.7080 |
| 6 | -52.4210 | -1.7600 | 6.5351 | 0.3475 | 2.6805 |
| 7 | -9.5920 | -1.2052 | -3.0648 | 0.1160 | 0.0752 |
| 8 | -79.2911 | -1.8617 | 7.3786 | 0.3402 | 1.7784 |
| 9 | -12.2207 | 0.2114 | 3.4196 | 0.1964 | 0.8877 |

| Matriz de correlación | | | | | |
|-----------------------|---------|---------|---------|--------|--------|
| | 1 | 2 | 3 | 4 | 5 |
| 3 | -0.6711 | 0.4452 | 1.0000 | 0.3853 | 0.4945 |
| 4 | -0.4001 | 0.4055 | 0.3853 | 1.0000 | 0.0599 |
| 5 | -0.5383 | -0.0256 | 0.4945 | 0.0599 | 1.0000 |
| 6 | -0.7370 | -0.3966 | 0.2038 | 0.2019 | 0.3560 |
| 7 | -0.2198 | -0.4427 | -0.1558 | 0.1099 | 0.0163 |
| 8 | -0.7468 | -0.2810 | 0.1542 | 0.1324 | 0.1582 |
| 9 | -0.5649 | 0.1566 | 0.3507 | 0.3752 | 0.3877 |

| V. propio | |
|-----------|--------|
| | 1 |
| 2 | 2.1302 |
| 3 | 1.0989 |
| 4 | 0.9945 |
| 5 | 0.5432 |
| 6 | 0.3834 |
| 7 | 0.2258 |
| 8 | 0.1368 |
| 9 | 0.0000 |

| Matriz de resultados | | | | | |
|----------------------|---------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | -0.1997 | -0.0715 | 0.0166 | -0.0040 | 0.0041 |
| 2 | -0.1253 | -0.2333 | -0.1457 | -0.1604 | 0.0160 |
| 3 | -0.1051 | -0.0918 | 0.0138 | -0.0040 | 0.0875 |
| 4 | -0.1568 | 0.0093 | 0.2461 | -0.0040 | -0.1031 |
| 5 | 0.0513 | -0.0513 | -0.1765 | 0.2045 | -0.0793 |
| 6 | -0.0408 | -0.1322 | 0.0166 | -0.2125 | 0.2186 |
| 7 | -0.1442 | 0.3732 | 0.1061 | -0.0561 | 0.1233 |

| Matriz de componentes | | | | | |
|-----------------------|---------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| 3 | -0.0000 | 0.0000 | -1.0989 | 0.0226 | -0.0342 |
| 4 | -0.0000 | 0.0000 | 0.0002 | -0.9945 | -0.0356 |
| 5 | -0.7209 | -0.0000 | -0.0000 | 0.0000 | -0.5432 |
| 6 | -0.9869 | -0.0996 | 0.0000 | 0.0000 | 0.0000 |
| 7 | -0.2944 | 0.3818 | 0.6616 | 0.0000 | -0.0000 |
| 8 | 0.0479 | 0.0458 | 0.2155 | -0.8195 | 0.0000 |
| 9 | -0.7565 | 0.2480 | -0.6923 | -0.6879 | -0.0804 |

| V. total | |
|----------|---------|
| | 1 |
| 2 | 23.6686 |
| 3 | 12.2102 |
| 4 | 11.0503 |
| 5 | 6.0358 |
| 6 | 4.2603 |
| 7 | 2.5084 |
| 8 | 1.5199 |
| 9 | 0.0005 |

Preparado NUM

Capítulo 7. Conclusiones

En el presente trabajo, se logró cubrir el objetivo del desarrollo de un sistema de software, enfocado a calcular las Componentes Principales de una matriz utilizando los algoritmos de Hessenberg y QR-Hessenberg, sistema donde se hizo operativa la aplicación de la ingeniería de software en su desarrollo, así como las aplicaciones de los Métodos Numéricos al Álgebra de Matrices y la Estadística, cuya operatividad se puso en práctica con empleo de un lenguaje de programación como es C++. Deja entre otras satisfacciones, el haber logrado que pueda ser fácilmente empleado por investigadores en Estomatología, su confianza se logra con las pruebas de cierta generalidad, para que sea utilizado por personas de áreas como la Biología y Economía entre otros.

Se logró llegar a resultados que probaron sus beneficios, al compararse con otros trabajos ya realizados con anterioridad, lo que le da robustez y precisión aceptable a los resultados obtenidos.

La elaboración de esquemas simples, como el uso de la creación de los diagramas que representan las actividades con las que se deseaba que contara el sistema, permite que al usuario final le sea de fácil entendimiento. Durante la etapa de diseño se codificaron e integraron todos los elementos que componen al sistema para que se implementara, se hicieron las pruebas pertinentes al mismo y se obtuvieron los resultados esperados al compararlos con los ejemplos indicados.

Si se considera que un sistema es aceptado al implementar su operatividad, pasando de las fases de prueba a las de productividad, y durante el uso que se haga de él, se pone en práctica también la etapa de mantenimiento conforme avanza el tiempo, por lo que se puede dar por concluida la presente tesis, dejando como posibles mejoras al sistema las que se mencionan a continuación:

- Uso de otros métodos para comparar los resultados obtenidos.
- Manejo de matrices de tamaño variable para procesar matrices de mayor tamaño.
- Capacidad de hacer un análisis de componentes principales.

Anexo A. Componentes principales

El análisis de componentes principales (ACP), es una técnica estadística propuesta a principios del siglo pasado por Karl Pearson como parte del análisis de factores. Sin embargo la complejidad de los cálculos retrasó su desarrollo hasta la aparición de los computadores y su utilización en la segunda mitad del siglo XX. Podría decirse que el objetivo principal que persigue el ACP es la representación de las medidas numéricas de varias variables en un espacio de pocas dimensiones donde nuestros sentidos puedan percibir relaciones que de otra manera permanecerían ocultas en dimensiones superiores. Para el caso de estudio de esta tesis, se tomaron solamente los cálculos de los valores propios de una matriz de covarianzas o de correlación como los resultados esperados.

Anexo B. Consideraciones sobre instalación y ejecución

La instalación del sistema desarrollado es realmente simple. Lo único que necesita es copiarlo a una carpeta (por ejemplo "c:\ccp") y asegurarse que en el directorio %windir%\system32 se encuentre el siguiente archivo (componente ocx):

- msflxgrd.ocx

En dado caso de que no se encuentre habrá que copiarlo (se distribuye junto con el programa) al directorio mencionado y ejecutar el comando siguiente desde el símbolo del sistema:

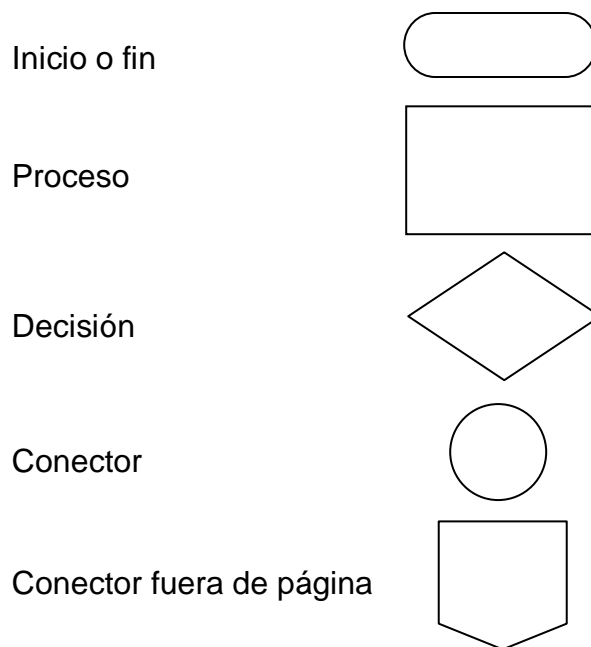
- regsvr32 msflxgrd.ocx

Esto registra el componente para que las aplicaciones de Windows lo puedan utilizar.

Una vez que se cuenta con los archivos mencionados, puede crear un acceso directo para ponerlo en el escritorio de Windows y ejecutarlo haciendo doble clic en él.

Anexo C. Diagramas de flujo

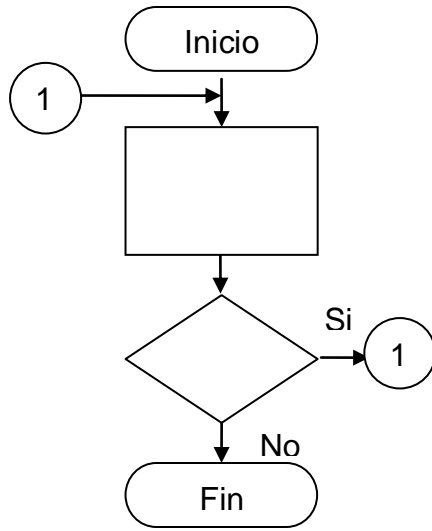
Los diagramas de flujo son ampliamente utilizados en el desarrollo de sistemas o inclusive en la documentación de cualquier procedimiento. Para esto se incluye una simbología especial que describe las actividades realizadas y el flujo del procedimiento inclusive a través de decisiones e iteraciones. Dentro de los símbolos más utilizados se encuentran:



Algunas de las convenciones utilizadas en los diagramas de flujo son:

- Debe tener un inicio y un fin.
- El flujo debe de ir de arriba hacia abajo y de izquierda a derecha.

Como ejemplo de un diagrama de flujo presentamos la siguiente figura.



Anexo D. Visual C++

Este ambiente de desarrollo proporciona una gran cantidad de elementos que ayudan a generar aplicaciones que trabajan con el ambiente del sistema operativo Windows. Entre las características que proporciona este ambiente para el desarrollo de aplicaciones podemos mencionar las siguientes:

- Facilidad de traducción del diseño al código.
- Portabilidad del código fuente.
- Facilidad de mantenimiento.

Por lo que respecta a los usuarios, obtendrán un sistema que se puede ejecutar en casi cualquier equipo y no se requiere mucho tiempo para aprender a utilizarlo, ya que es muy similar a cualquier otro sistema con la misma interfaz estándar de los sistemas en Windows.

Los componentes principales que Visual C++ proporciona son:

- Editor de texto.
- Editor de recursos.
- Facilidades para construcción de proyectos.
- Depurador integrado

La parte más importante de Visual C++ es un asistente llamado ClassWizard, el cuál es nos permite indicarle dónde queremos escribir el código para las actividades asociadas al aspecto visual del sistema.

Bibliografía

- [1] Roger S. Presman. Ingeniería de Software, un enfoque práctico. 4ª. Edición. Mc Graw Hill.
- [2] Richard L. Burden, J. Douglas Faires. Análisis Numérico. 6ª. Edición, Youngstown State University. International Thomson Editores.
- [3] J. M. Caldera. Identificación de cambios físicos en el esmalte dental utilizando el microcontrolador 8032, tesis de licenciatura, BUAP, 2003.
- [4] Bryan F. J. Manly. Multivariate Statistical Methods. Department of Mathematics and Statistics University of Otago, New Zealand. 1986. Chapman and Hall.
- [5] Chris H. Pappas, William H. Murray III. Visual C++ 6: the complete reference. Osborne Mc Graw Hill.
- [6] I.T. Jolliffe (1986): *Principal Component Analysis*. Ed. Springer-Verlag, New York.