

SISTEMA INTELIGENTE PARA EL ALGEBRA LINEAL

Laura Casas Fuentes, Olga Lidia Pérez González, Lisandra Docampo, Yailé Caballero Mota, Lenniet Coello, Isabel Yordi González, Angela Martín.

Facultad de Informática, Universidad de Camagüey
Universidad APEC

Cuba
República Dominicana

laura.casas@reduc.edu.cu, lisandra.docampo@reduc.edu.cu, olga.perez@reduc.edu.cu; m.angela24@gmail.com

Resumen. El trabajo describe la concepción del Sistema Inteligente para el Álgebra Lineal (SIAL), el cual fue creado por los autores de la investigación. Sus fundamentos teóricos están relacionados con las técnicas de Inteligencia Artificial para clasificar, utilizando el clasificador k-vecinos más cercanos K-NN (Caballero, 2010), así como los cinco problemas tipos del Álgebra Lineal (Yordi, 2004). El SIAL se creó con el objetivo de apoyar al estudiante durante su estudio independiente para brindarle la vía de solución de los problemas que debe resolver.

Palabras clave: sistema inteligente, algebra lineal, inteligencia artificial

Abstract. This work describes the design of the Intelligent System for Linear Algebra (SIAL); it was created by the authors of this research. Its theoretical foundations are related to Artificial Intelligence techniques for classifying, using the k-nearest neighbor K-NN classifier (Caballero, 2010), and the five types of Linear Algebra problems (Yordi, 2004). The SIAL was created to help the student during their independent study, this provides the way to solve the problems oriented in the assignment.

Key words: intelligent system, linear algebra, artificial intelligence

Introducción

Las investigaciones pedagógicas relacionadas con la aplicación de la computadora en la enseñanza de la Matemática se centran fundamentalmente en la introducción de herramientas (software), que faciliten el proceso de cálculo. La diversidad de software que anualmente se genera en la esfera de la Enseñanza Asistida por Computadora es muy amplia. No obstante, resulta difícil encontrar software educativos que se ajusten a los requerimientos de un proceso activo de aprendizaje.

Se hizo un análisis de los asistentes matemáticos que, por lo general, facilitan el cálculo de operaciones de Geometría Analítica y Álgebra Lineal: el Derive, MATLAB; mientras otros como la plataforma interactiva MOODLE, con actividades evaluables, lo que hace es procesar información, pero no está orientada a la búsqueda del conocimiento, por lo tanto se hace necesaria la sustitución del producto "información" por el producto "conocimiento" y de "sistemas que permiten procesar información" por sistemas que generan o entregan conocimientos, es decir que aseguren el uso productivo de la información, que guíen una toma de decisión óptima, capaces de tomar decisiones inteligentes en un determinado tema, con la capacidad de adquirir nuevos conocimientos y perfeccionar el que posee.

Para el desarrollo de herramientas que superen estas dificultades, las técnicas de Inteligencia Artificial resultan de mucho interés, debido a todos los métodos desarrollados para la adquisición del conocimiento y el aprendizaje automático, el cual consiste en generalizar comportamientos a partir de una información no estructurada suministrada en forma de ejemplos. Esto sugirió un estudio teórico-práctico de las relaciones que se puedan establecer entre la Inteligencia Artificial y la autoevaluación del estudiante; de forma que con el empleo de las técnicas de aprendizaje se puedan desarrollar herramientas de apoyo a la actividad independiente que realizan los estudiantes y al sistema de ayudas pedagógicas que responda a las especificidades de cada asignatura.

En este contexto, la Inteligencia Artificial puede permitir fundamentar una nueva línea de trabajo, orientada a diseñar programas de consulta, capaces de ayudar a resolver dudas al estudiante, a los que se les da el nombre de "Sistemas Expertos" (SE). Son programas que reproducen el proceso intelectual de un experto humano en un campo particular, pudiendo mejorar su productividad, ahorrar tiempo y dinero, conservar sus valiosos conocimientos y difundirlos más fácilmente. Además son capaces de arribar a conclusiones a partir de la información almacenada en su base de conocimiento.

Dado lo abstracto del contenido de Espacio Vectoriales (EV) en Álgebra Lineal (AL), el estudiante requiere de un tiempo prolongado de estudio independiente, pues no asimilan las definiciones, conceptos, por lo general los ejercicios tienen varias vías de solución y los alumnos no son capaces de seleccionar el algoritmo correcto de los propuestos por el profesor. Una vez identificado que las herramientas informáticas disponibles no les resultaban factibles para analizar las posibles vías de solución de una tarea dada, se defiende la idea de la utilización de Sistemas Expertos que propongan caminos de solución a los problemas y no la solución en sí.

Metodología

En consideración a los aspectos anteriores, la investigación tuvo como objetivo el de implementar un SIAL, aplicando las técnicas de Inteligencia Artificial, con el fin de que el estudiante al realizar una tarea, introduzca los datos e incógnita de la misma y el software le ofrezca la vía de solución a la tarea y lo oriente qué debe estudiar para poder comprender la tarea, por lo que se debía disponer de un módulo inteligente.

En el artículo se describe el proceso seguido para la producción de un software, que se fundamenta en cinco problemas tipos para el Álgebra Lineal, en técnica la Inteligencia Artificial, para clasificar, utilizando el clasificador k-vecinos más cercanos K-NN y en la herramienta de la inteligencia artificial llamada razonamiento basado en casos.

El software desarrollado es catalogado como un Sistema de Experto. Se describen detalladamente las 7 acciones que realizaron sus autores desde la descripción de los fundamentos hasta el diseño de la interfaz de usuario. Al finalizar se hace una breve descripción del impacto que se ha tenido con su implementación. Las 7 acciones desarrolladas constituyen una guía para desarrollar otros sistemas de expertos para las Matemáticas, siendo este el principal aporte del trabajo. Se asume como innovación al proceso de producción del Sistema De Experto.

Fundamentos teóricos

Los cinco tipos de problemas de Álgebra Lineal. (Yordi, 2004)

Desde el punto de vista didáctico, se ha trabajado con la estructuración sistémica de la asignatura, para favorecer la enseñanza-aprendizaje de esta asignatura, en este sentido se han determinado cinco problemas tipos que relacionan el contenido entre los temas y el contenido dentro del tema. La clasificación en esos tipos de problemas favorecen la estructuración sistémica de la asignatura y permiten orientar al estudiante en la ejecución de las tareas, en su actividad de estudio independiente.

Muchos de los principales problemas con la Combinación Lineal de vectores se reducen a los problemas tipos con las formas lineales (Yordi, 2004). Una forma lineal es una combinación lineal, de allí que los principales problemas del AL puedan enunciarse en términos de la teoría de Espacios Vectoriales.

Dado el sistema de ecuaciones en la forma matricial $AX=Y$, si se denotan los vectores columna de la matriz A como $a_i, i = 1 \dots n$, los problemas tipo quedarían enunciados de la siguiente forma

- ❖ Problema directo: Dados los vectores de la Combinación Lineal a_i y los coeficientes de dicha combinación lineal x_i (elementos del vector X), calcular el vector resultante de dicha Combinación Lineal Y (de componentes y_i).
- ❖ Problema inverso: Dados los vectores de la Combinación Lineal a_i y el vector resultante de dicha Combinación Lineal Y (de componentes y_i). Calcular los coeficientes de dicha Combinación Lineal x_i (elementos del vector X).
- ❖ Problema de dependencia de los datos: Dados los vectores a_i y el vector Y (de componentes y_i). Plantear la interrogante de la dependencia lineal de estos vectores (existencia de los x_i , únicos o infinitos) o la independencia lineal de los mismos (no existencia de los x_i).
- ❖ Problema de factibilidad: Dados los vectores a_i y el vector Y (de componentes y_i). Cuestionar la posibilidad de generar el vector Y , mediante la combinación lineal de

los vectores a_i , siendo los coeficientes x_i (coeficientes de la combinación lineal) únicos, infinitos o si no existen (no es posible generar el vector Y combinando linealmente los vectores a_i).

- ❖ Problema de unicidad: Dados los vectores a_i y el vector Y (de componentes y_i). Cuestionar la posibilidad de generar el vector Y , mediante la combinación lineal de los vectores a_i , siendo únicos los coeficientes x_i (coeficientes de la Combinación Lineal). No obstante, a pesar de lo mucho que se ha avanzado en la enseñanza del Álgebra Lineal aún existen insuficiencias, se requiere de un perfeccionamiento dentro de la disciplina Matemática para carreras técnicas.

Las técnicas de Inteligencia Artificial para clasificar, utilizando el clasificador k-vecinos más cercanos K-NN. (Caballero, 2010)

La clasificación consiste en el proceso de asignar a una entrada concreta, el nombre de una clase a la que pertenece. Las clases se pueden describir de gran cantidad de formas. Su definición dependerá del uso que se les dé.

La clasificación constituye una parte importante de muchas de las tareas de resolución de problemas. En su forma más simple, se presenta directamente como una tarea de reconocimiento de un objeto a determinado clase. Antes de que se pueda hacer la clasificación, se deben definir las clases que utilizará. (Caballero, 2010). Los diferentes algoritmos de aprendizaje se agrupan en una taxonomía en función de la salida de los mismos. Algunos tipos de algoritmos son:

- ❖ *Aprendizaje supervisado*. se conocen las entradas y las salidas deseadas del sistema. Un ejemplo de este tipo de algoritmo es el problema de clasificación, donde el sistema de aprendizaje trata de etiquetar (clasificar) una serie de vectores utilizando una entre varias categorías (clases). La base de conocimiento del sistema está formada por ejemplos de etiquetados anteriores.
- ❖ *Aprendizaje no supervisado*. todo el proceso de modelado se lleva a cabo sobre un conjunto de ejemplos formado tan sólo por entradas al sistema. No se tiene información sobre las categorías de esos ejemplos

Razonamiento basado en casos

El Razonamiento Basado en Casos (RBC) es una de las múltiples herramientas con las que cuenta hoy la Inteligencia Artificial (IA). Con esta herramienta la solución de un nuevo problema se realiza a partir de las soluciones conocidas para un conjunto de problemas previamente resueltos (o no resueltos) del dominio de aplicación. Su rasgo distintivo es el hecho de utilizar

directamente la información almacenada en la memoria del sistema sobre los problemas o casos. Los sistemas que emplean el RBC usan una memoria permanente en lugar de alguna forma de base de conocimientos en la cual se almacene de forma explícita el conocimiento sobre el dominio de aplicación, en forma de estructuras conceptuales, reglas de producción u otra forma de representación del conocimiento.

RBC significa razonar en base a experiencias. Es una alternativa entre otras metodologías para construir sistemas basados en el conocimiento que se asemejan en gran medida a la forma de razonamiento humano. Al razonar basado en casos el solucionador de problemas recuerda situaciones previas similares a la actual y las usa para ayudar a resolver el nuevo problema.

k-Vecinos más Cercanos (k-NN)

El clasificador *k-Vecinos más Cercanos (k-NN)* es uno de los métodos de clasificación supervisada más usados que se usa en el razonamiento basado en casos. Utiliza funciones de distancia o similitud para generar predicciones a partir de ejemplos almacenados.

Este algoritmo pertenece a la clase de algoritmo perezoso (*Lazy Learning algorithms*), pues almacena el conjunto de entrenamiento y deja todo el procesamiento para la fase de clasificación. Es altamente sensible a la definición de las funciones de distancia o similitud utilizadas y a las características de los rasgos que conforman cada uno de los ejemplos almacenados. Muchas variantes del *k-NN* se han propuesto reducir esta sensibilidad parametrizando la función de distancia o similitud con características de peso. Otro de los factores que intervienen en el desempeño del clasificador es la base de casos que se utiliza en el proceso de clasificación.

La entrada del clasificador es una instancia q , cuya clase se desconoce, y la salida es su clasificación. Los rasgos pueden ser continuos o pueden estar limitados dentro de un conjunto fijo de valores discretos. El error que puede cometerse al clasificar cada instancia del conjunto de entrenamiento es conocido como *Error de Clasificación Dejando uno Fuera (LOOCE* por sus siglas en inglés). El objetivo del clasificador es minimizar el coeficiente LOOCE y existen algoritmos eficientes para minimizarlo. La forma de calcular el LOOCE está en dependencia de si los valores de la clase son continuos o discretos. Para valores discretos se calcula como:

$$(1) \text{ LOOCE} = \sum_{q \in BC} \sum_{j \in J} (\delta_{q,j} - p_{q,j})^2$$

Donde q es cada instancia que pertenece a la base de casos BC y j cada clase que pertenece al conjunto de clases J . La función de probabilidad de clase $p_{q,j}$, definidas como:

$$(2) \delta_{q,j} = \begin{cases} 1 & q_c = j \\ 0 & q_c \neq j \end{cases}$$

$$(3) p_{q,j} = \frac{\sum_{r \in K} \delta_{r,j} \cdot \text{sim}(q,r)^2}{\sum_{r \in K} \text{sim}(q,r)^2}$$

Donde K es el conjunto de las instancias vecinas más similares

Las acciones científico técnicas que dieron lugar a la innovación fueron las siguientes (Caballero, Pérez, Docampo, Casas, & Yordi, 2011)

En el artículo se describe el proceso seguido para la producción de un software, que se fundamenta en las técnicas de Inteligencia Artificial, a través de siete acciones que se describen detalladamente, las cuales comienzan con la fundamentación de la propuesta hasta el diseño de la interfaz de usuario. Al finalizar se hace una breve descripción de la experiencia que se ha tenido con su implementación. La metodología seguida para desarrollar el trabajo es descriptiva y su importancia radica en mostrar los pasos que han de seguirse para la producción de tecnología educativa utilizando dichas técnicas. Se asume como innovación al proceso de producción de dicha tecnología.

La *primera acción*, la cual es un fundamento científico de la innovación tecnológica, estuvo dirigida a caracterizar el proceso de enseñanza aprendizaje del Álgebra Lineal en las carreras de Ingeniería. En la *segunda acción* se trabajó en los sustentos teóricos, descrito anteriormente, precisándose el Razonamiento Basado en Casos (búsqueda por semejanza y adaptación de las soluciones). Los experimentos realizados con herramientas como Weka, MLP, KNN-WorkShop; determinaron que el clasificador k-vecinos más cercanos (K-NN) es el óptimo para diseñar e implementar la máquina de deducción del SIAL, la cual proporciona la capacidad de razonamiento del SE, al extraer conclusiones sobre la clasificación.

En la *tercera acción* se diseña e implementa el módulo inteligente o módulo “Obtener vía de solución para el tema de Espacios Vectoriales del Álgebra Lineal”, usando la base de conocimiento diseñada y el clasificador para inferir soluciones. A partir de encuestas, entrevistas con expertos, se determinaron los requisitos funcionales relacionados con orientar al estudiante hacia la búsqueda de una solución a un problema dado, tomando como base la clasificación del mismo en uno de los cinco problemas tipo del Álgebra Lineal, de manera que el SIAL no sólo oriente al estudiante, sino que le permita justificar los resultados de una clasificación. Los requisitos funcionales definidos fueron: ofrecer la vía de solución a un problema, clasificar en un problema tipo los datos introducidos por el estudiante, mostrar los resultados de la clasificación, justificar los resultados de la clasificación, guardar los resultados

de la clasificación en la base de datos y asignar entrenador (material de consulta al estudiante con dificultades). Luego de realizar un estudio de las características del negocio y el problema a automatizar se decidió que estos requisitos conformarían el módulo “Obtener vía de solución del sistema”. Esta opción hace del SIAL un sistema computacional con características especiales, ya que incorpora en forma operativa el conocimiento de los profesores más experimentados en la asignatura Álgebra Lineal, de forma que es capaz de resolver problemas inteligentemente y de justificar sus respuestas.

En la *cuarta acción* se diseñó e implementó la base de conocimiento. Esta es el componente más importante de un Sistema Experto y tiene asociado un formato el cual indica cómo el conocimiento será representado. La base de conocimiento del SIAL proporciona ejemplos sobre la asignatura que es objeto de estudio y está basada en el razonamiento en casos, con 22 rasgos, 6 clases y 66 instancias. Se determinó que un caso estará compuesto por un vector de datos y su clasificación. Los elementos que forman el vector van a ser los datos seleccionados por el estudiante como la información que tiene del problema y las incógnitas. La clasificación va a ser clase de los problemas absurdos o uno de los cinco problemas de la combinación lineal definidos por los especialistas.

A continuación se muestran los 22 rasgos, del 1 – 12 representan los posibles datos de un problema planteado por el alumno y del 13 – 22 la posibles datos que necesita conocer

1. Vectores de R_n
2. Sistema de matrices
3. Sistema de vectores
4. Escalares de la CL
5. Vector resultante/ matriz resultante
6. Base del EV
7. Vector genérico
8. EV (condiciones)
9. Sistema de vectores canónicos/ Matriz idéntica
10. Sistema generador
11. Linealmente Independiente (LI)
12. Linealmente Dependiente (LD)
13. Vector resultante
14. Escalares de la CL
15. Coordenadas

16. LD
17. LI
18. Expresar un vector como CL de los demás
19. Base
20. Vectores resultantes (vector genérico). Subespacio generado
21. Número máximo de vectores LI/Dimensión
22. Polinomio

En la *quinta acción* selección del Razonamiento Basados en Casos como método para la adquisición del conocimiento y construir sistemas basados en el conocimiento. Su principal ventaja es que este método tiene la facilidad de que el conocimiento que se almacena puede ser ampliado y por consiguiente, la exactitud de la clasificación, así como la veracidad de las posibles respuestas aumenta. Es muy difícil considerar en una base de casos todas las selecciones que puede realizar un estudiante, algunas, en un número elevado de casos, carentes de coherencia y sentido lógico y totalmente alejadas de los conceptos de Álgebra Lineal. Es por este motivo que se decidió incluir una sexta clase que agrupa estos últimos casos. Luego se comenzaron a adicionar casos a la base para lo cual fue necesario realizar varias consultas a expertos. Se ofrece una base de casos de prueba con los 22 rasgos mencionados anteriormente, 1 rasgo de clase con un dominio de 1-6 y 66 instancias o problemas ejemplos que han sido ejecutados exitosamente sobre la base de conocimiento. Finalmente, se definió que un caso estará compuesto por un vector de datos y una clasificación. Los elementos que forman el vector van a ser los datos seleccionados por el estudiante como la información que tiene del problema y las incógnitas. La clasificación va a ser uno de los cinco problemas de la combinación lineal definidos por los especialistas o la clase de los problemas absurdos.

En la *sexta acción* se definió el mecanismo de explicación, el cual consiste en mostrar al usuario cómo se llegó a la clasificación del nuevo caso. La forma de dar la conclusión es variada y en este caso se realizó de la siguiente forma: lo primero es clasificar, a través del kNN, el nuevo objeto en una de las 6 clases. Para cada clase estará registrada en la base de datos su correspondiente explicación de cómo se resuelve ese tipo de problema. Debido a que $k=1$, se obtiene el vecino más cercano, que es, de los objetos registrados en la base de casos, el más similar al nuevo objeto. De sus rasgos se presentarán solo aquellos con valor 1, es decir, los que indican una posible selección, los cuales serán mostrados en el reporte.

En la *séptima acción* se diseñó la interfaz de usuario. Mediante la interfaz el estudiante plantea los problemas al sistema y este ofrece la vía de solución, o sea, la clasificación en uno de los 5 problemas tipos. Al seleccionar la opción obtener vía de solución se le mostrará una nueva ventana para que seleccione los datos de su problema y las incógnitas, los cuales serán identificados por el sistema como atributos de la BC. El sistema derivará una clasificación y mostrará al estudiante por cuál vía puede resolver su ejercicio. Cada clase se vincula con varios entrenadores, asignados aleatoriamente evitando obtener el mismo material para clasificaciones iguales. El entrenador servirá para que el estudiante interiorice los conceptos y definiciones de la teoría de Espacios Vectoriales implícitas en su problema. Los mismos estarán relacionados con uno de los cinco problemas de la Combinación Lineal. De esta forma el estudiante obtendrá una vía de solución a su ejercicio y así podrá resolverlo por diferentes caminos.

Implementación en la práctica (Pérez, Caballero, Docampo, Casas, & Yordi, 2012)

Desde el curso 2008-2009 se está implementando el SIAL en los cursos de Álgebra Lineal en la Universidad de Camagüey, Cuba. A través de su uso se valoró si realmente el sistema es utilizado por los estudiantes como apoyo durante su estudio independiente y si ellos consideraban que el SIAL les brindaba la vía de solución de los problemas que debe resolver. En todos los cursos se aplicó una encuesta dirigida a conocer las opiniones de los estudiantes y los resultados fueron favorables.

Una vez que se adquiere el sistema puede ser duplicado lo cual reduce sus costos. Al poder obtener información de una base de conocimiento puede realizar cálculos y derivar soluciones mucho más rápido que cualquier experto humano. El sistema puede simular el rol de un profesor de Álgebra Lineal lo cual lo convierte en un apoyo académico muy valioso para los estudiantes de los cursos semipresenciales. El sistema contribuirá en gran medida a crear habilidades en los estudiantes y consolidar conocimientos durante el estudio independiente. Además permitirá a los profesores conocer y tratar las dificultades de sus estudiantes durante el proceso de enseñanza-aprendizaje para así crear mecanismos que tributen a elevar la calidad del proceso.

El impacto científico del resultado se evidencia en la obtención del Premio Territorial de la Academia de Ciencias de Cuba (2010) y el premio del Rector al resultado, ya aplicado, de mayor impacto al desarrollo social (2011). Actualmente se desarrollan SIAL para los temas de Diagonalización y Aplicaciones Lineales.

Referencias bibliográficas

- Caballero, Y., Pérez, O., Docampo, L., Casas, L., Yordi, I., Coello, y otros. (2011). Sistema Experto para el Álgebra Lineal. *XII Congreso de la Sociedad Cubana de Matemática y Computación (COMPUMAT2011)*, (págs. 25-37). Villa Clara, Cuba.
- Caballero, Y. (2010). La Teoría de los Conjuntos Aproximados para el Descubrimiento de Conocimiento. *Revista DYNA*, 77 (162), 261-270.
- Pérez, O., Caballero, Y., Docampo, L., Casas, L., Yordi, I., Coello, y otros. (2012). Valoración del impacto de la implementación del SIAL. *Informe de Ciencia y Técnica*, (pp. 5-21). Vicerrectoría de Investigaciones, Universidad de Camagüey, Cuba.
- Yordi, I. (2004). *Metodología para formar en los estudiantes de Ingeniería Eléctrica la habilidad de calcular en Álgebra Lineal con sentido amplio*. Tesis doctoral no publicada, Universidad de Camagüey, Departamento de Matemática, Camagüey, Cuba.