Presentació

Aquesta PAC és una introducció a conceptes com espais i subespais vectorials, combinació lineal, independència lineal, dimensió, matrius i vectors, determinants, i equacions de rectes i plans en l'espai que seran molt útils per resoldre problemes típics de la ciència de dades. Alguns conceptes de la geometria mètrica com el producte escalar, la ortonormalitat, els angles i les distàncies també seran tractats en aquesta PAC.

Competències

En aquesta PAC es treballen les següents competències del Grau en Ciència de Dades Aplicada:

- Que els estudiants hagin demostrat tenir i comprendre coneixements en un àrea d'estudi que parteix de la base de l'educació secundària general, i se sol trobar a un nivell que, si bé es recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.
- Utilitzar de forma combinada els fonaments matemàtics, estadístics i de programació per desenvolupar solucions a problemes en l'àmbit de la ciència de les dades.
- Ús i aplicació de les TIC en l'àmbit acadèmic i professional.

Objectius

Els objectius concrets d'aquesta PAC són:

- Comprendre la importància de l'àlgebra lineal en l'àmbit de la ciència de dades.
- Conèixer i ser capaç de manipular elements bàsics d'àlgebra lineal (espais vectorials, independència lineal, dimensió, matrius, determinants) i de la geometria mètrica (productes escalars, ortonormalitat, angles i distàncies).
- Ser capaç d'utilitzar software matemàtic per a resoldre problemes bàsics de l'àlgebra lineal.

Descripció de la PAC

Aquesta activitat ens permetrà posar en pràctica els coneixements i procediments treballats en aquest repte.

Per una banda, us demanem que respongueu un qüestionari (el podeu trobar al moodle entrant a l'enllaç "Qüestionaris"a la part dreta de l'aula) en el què treballarem la part més instrumental d'aquest repte en una sèrie de preguntes genèriques. Podeu practicar amb el qüestionari "Repte 2 entrenament"els intents que creieu necessaris (els resultats no es tindran en compte per la nota). Un cop us sentiu segurs podeu resoldre el qüestionari "Repte 2 avaluació", penseu que per aquest només teniu 5 intents. Llegiu el FAQ que trobareu a l'enllaç de "Qüestionaris".

Us demanem també que resolgueu una sèrie d'exercicis on s'hauran d'aplicar els procediments que hem anat treballant en el context de la ciència de dades. Aquests exercicis us plantejaran escenaris propis de la ciència de dades i veurem com els conceptes treballats en aquest repte tenen rellevància en aquests contexts.

Recursos

Recursos Bàsics

- Document introductori als elements bàsics de l'àlgebra lineal per a la ciència de dades
- Mòdul Elements d'àlgebra lineal i geometria
- Document de problemes sobre elements bàsics de l'àlgebra lineal enfocats a la ciència de dades
- Vídeo resum: Espais vectorials

Recursos Complementaris

- Calculadora CalcMe
- Documentació de CalcMe
- Guia CalcMe UOC

- Laboratori CalcMe

Criteris d'avaluació

- LA PAC s'ha de resoldre de manera individual.
- És necessari justificar totes les respostes proposades al lliurable de la PAC.

Tingueu en compte que les dues activitats que es plantegen a aquest repte (la resolució dels exercicis que es plantegen d'aquest document i el qüestionari) seran part de la nota d'avaluació contínua AC: (PAC1 + PAC2 + PAC3)/3. La nota d'aquestes activitats correspon a la PAC1 (amb un pes del 70% pel qüestionari i un 30% pel lliurable). **Recordeu que heu d'aprovar l'AC per superar l'assignatura.** Per a més informació sobre el model d'avaluació de l'assignatura consulteu el pla docent.

Format i data de lliurament

Cal lliurar un **únic document PDF** amb les respostes de tots els exercicis. Cal realitzar la PAC **amb un processador de textos**, no s'acceptaran solucions a mà i escanejades.

El nom del fitxer ha de ser PAC2Cognom1Cognom2Nom.pdf. Aquest document s'ha de lliurar a l'espai de Lliurament i Registre AC de l'aula abans de les 23:59 del dia 16/04/2021. No s'acceptaran lliuraments fora de termini.

Què veus en observar la imatge de dalt? Sens dubte, roses de diferents colors. Ara bé, una tasca difícil seria escriure aquesta lògica per a que un ordinador pogués deduir-ho tan fàcilment com nosaltres.

De manera senzilla heu identificat la flor perquè el cervell humà ha evolucionat al llarg de milions d'anys. No importa si els colors de les flors no són els habituals. Hem entrenat d'alguna manera els nostres cervells per realitzar aquesta tasca automàticament.

Però fer que un ordinador faci la mateixa tasca no és una tasca fàcil. La ciència de dades, i l'aprenentatge automàtic se n'ocupen de buscar solucions a aquest problema. En primer lloc cal identificar els atributs de la imatge i representar-los. Ens plantegem, en aquest cas, com s'emmagatzema el color en un ordinador.

Ja al segle XIX, els científics Thomas Young i Hermann von Helmholtz van observar que els colors visibles poden generar-se mitjançant barreges de tres feixos de llum de colors diferents, que formen una *base* de colors. Els colors usats a la base poden variar. Una elecció comú és utilitzar llum de colors vermell, verd i blau. Per les seves sigles en anglès usarem les lletres R (*red:* vermell), G (*green:* verd) i B (*blue:* blau) per indicar llum dels colors de la base. En barrejar els feixos de llum, es genera un nou color que correspon a una combinació lineal dels colors de la base:

$$C = rR + gG + bB,$$

on r, g, b són les potències de les fonts de llum vermella, verda i blava respectivament.

Això permet representar el color C per la tripla (r, g, b) dels coeficients de la combinació lineal dels tres colors de la base. Per exemple: $(1, 0, 0)$ és el color vermell, $(0, 0, 0)$ és negre (absència de llum) i $(1, 1, 1)$ genera el blanc més clar permès per la potència de les fonts.

El problema és que a la pràctica, és molt difícil fabricar fonts de llum exactament iguals. Aquí és on l'àlgebra lineal és de gran ajuda, oferint, per exemple per aquest cas, canvis de base.

Les matrius, per exemple, ens permeten representar una imatge. Així, podem descompondre una imatge determinada en els tres canals de color vermell, verd i blau. I cada canal es pot representar com una matriu $m \times n$ amb valors que oscil·len entre 0 i 255. Estudiar les propietats de les matrius és molt rellevant. Veiem tot seguit algunes.

Considereu una matriu A amb m files i n columnes. Definim el conjunt $N(A)$ com

$$N(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}.$$

1. Proveu que $N(A)$ és un subespai vectorial.

Solució:

Per a veure que és un subespai vectorial, comprobarem que es compleixen les 3 propietats que ha de satisfer un subespai vectorial i que trobareu a la secció 2.2 del mòdul *Elements d'àlgebra lineal i geometria*.

1. Observeu que el 0 és una solució del sistema homogeni. Per tant, $0 \in N(A) \subset \mathbb{R}^n$.
2. Donats qualssevol $u, v \in N(A)$, anem a veure que $u + v \in N(A)$. En efecte, com $u, v \in N(A)$, $Au = 0$ i $Av = 0$. Per altra banda, donat que $A(u+v) = Au + Av = 0+0 = 0$ i $u+v \in N(A)$.
3. Donats $u \in N(A)$ i $k \in \mathbb{R}$, veurem que $ku \in N(A)$. Com $u \in N(A)$, sabem que $Au = 0$. Donat que $Akv = kAv = 0$, $kv \in N(A)$.

---

2. Considereu la matriu $A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 2 & 0 & 1 & -1 \\ 0 & 0 & 3 & 1 \end{pmatrix}$ i calculeu el complement ortogonal de $N(A)$.

Solució:

Seguint la definició de la Secció 6.2 del complement ortogonal, sabem que és el conjunt de tots els vectors u d' $\mathbb{R}^n$ que són ortogonals al subespai $N(A)$, per tant en primer lloc calcularem una base de $N(A)$ i a continuació calcularem el complement ortogonal.

Per a trobar una base de $N(A)$, necessitem resoldre el sistema homogeni $Ax = 0$. Ho farem utilitzant eliminació de Gauss:

$$\begin{pmatrix} 1 & 0 & -1 & 1 \\ 2 & 0 & 1 & -1 \\ 0 & 0 & 3 & 1 \end{pmatrix} \overset{F_2-2F_1}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 3 & 1 \end{pmatrix} \overset{F_3-F_2}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Per tant, la solució del sistema lineal és l'espai generat pel vector $(0, 1, 0, 0)$, és a dir, $N(A) =< (0, 1, 0, 0) >$.

Busquem ara el complement ortogonal a aquest espai $N(A)$, és a dir $N(A)^{\perp}$:

$$N(A)^{\perp} \{x = (x_1, x_2, x_3, x_4) \mid x \cdot (0, 1, 0, 0) = 0\} = \{x = (x_1, x_2, x_3, x_4) \mid x_2 = 0\} =$$

$$= \langle (1, 0, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1) \rangle.$$