Föreläsning 11

Innehåll: Huvudsatsen med tillämpningar

Kapitel 9.5–9.6

Huvudsatsen

Idag ska vi först samla den viktiga teorin för linjära ekvationssystem i den s.k. huvudsatsen:

Sats *För en kvadratisk matris är följande villkor ekvivalenta:*

1. *A:s kolonner utgör en bas,*
2. *A:s rader utgör en bas,*
3. *ekvationssystemet $AX = 0$ har endast den triviala lösningen $X = 0$,*
4. *ekvationssystemet $AX = Y$ är lösbart för alla Y,*
5. *A är inverterbar,*
6. *den linjära avbildning vars avbildningsmatris är A är bijektiv,*
7. $\det A \neq 0$.

Gå igenom gamla föreläsningar/boken och förvissa dig om när vi stötte på de olika påståendena och varför vi har ekvivalenser mellan dem. Ingenting är nytt! Men en del kan vara flerstegsresonemang.

Sambandet mellan determinanter och linjära ekvationssystem kan också förtydligas i följande tabell

	homogent system $AX = 0$	inhomogent system $AX = Y,\ Y \neq 0$
$\det A = 0$	oändligt många lösningar	ingen eller oändligt många lösningar
$\det A \neq 0$	endast den triviala lösningen	precis en lösning

Exempel För vilka värden på a har systemet

$$\begin{cases} ax + y + 2z &= 4 \\ x + y + z &= 1 \\ x + ay + 2z &= 0 \end{cases}$$

oändligt många lösningar.

Enligt satsen ovan ser vi att detta är fallet kan bara hända då $\det A = 0$:

$$\begin{vmatrix} a & 1 & 2 \\ 1 & 1 & 1 \\ 1 & a & 2 \end{vmatrix} = \begin{vmatrix} a-2 & 1 & 2 \\ 0 & 1 & 1 \\ -1 & a & 2 \end{vmatrix} = (a-2)\begin{vmatrix} 1 & 1 \\ a & 2 \end{vmatrix} + (-1)\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix}$$

$$1 - (a-2)^2 = (1-(a-2))(1+(a-2)) = (3-a)(a-1).$$

Vi ser att determinanten är noll precis så $a = 1, 3$. I alla övriga fall finns en entydigt bestämd lösning, men i dessa kan det antingen saknas lösning eller så finns det oändligt många. Vilket måste kontrolleras i varje fall:

1. I fallet $a = 1$ får vi systemet

$$\begin{cases} x + y + 2z &= 4 \\ x + y + z &= 1 \\ x + y + 2z &= 0 \end{cases} \Leftrightarrow \begin{cases} x + y + 2z &= 4 \\ z &= 3 \\ 0 &= 4 \end{cases},$$

   så i detta fall saknas lösning
2. I fallet $a = 3$ får vi systemet

$$\begin{cases} 3x + y + 2z &= 4 \\ x + y + z &= 1 \\ x + 3y + 2z &= 0 \end{cases} \Leftrightarrow \begin{cases} 3x + y + 2z &= 4 \\ -2y - z &= 1 \\ 0 &= 0 \end{cases},$$

   vilket betyder att lösningen ges av linjen $(x, y, z) = (2 + t, t, -1 - 2t)$. I detta fall finns alltså oändligt många lösningar.

Svaret är alltså $a = 3$.

Egenarbete

Gör uppgift **1.18**.

Ett kemi-exempel

Finns det någon reaktion som oxiderar metan (CH_4) till koldioxid (CO_2) och vatten (H_2O)? Att oxideras innebär att metanet ska reagera med syrgas (O_2). Vi söker alltså en ekvation av typen

$$x_1 CH_4 + x_2 O_2 \to x_3 CO_2 + x_4 H_2O$$

där x_i:na är heltal.

För att ta reda på detta måste vi ha massbalans i den meningen att reaktionen har lika många C, H, O-atomer före som efter. Det innebär att vi måste ha följande samband:

$$\begin{cases} C & x_1 = x_3 \\ H & 4x_1 = 2x_4 \\ O & 2x_2 = 2x_3 + x_4 \end{cases}.$$

Vi kan skriva detta som matrisekvationen

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 4 & 0 & 0 & 2 \\ 0 & 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} -x_1 \\ -x_2 \\ x_3 \\ x_4 \end{pmatrix} = 0.$$

Lägg märke till minutecknens placering! Vi sätter minustecken på de molekyler som finns till vänster om reaktionspilen. Kolonnerna i matrisen representerar den kemiska formeln för varje molekyl i sina beståndsdelar (C, H, O). T.ex. innebär första kolonnen ämnet $C_1H_4O_0 = CH_4$.

Eftersom vi har tre ekvationer men fyra obekanta vet vi att det finns lösningar på detta problem (dimensionssatsen sa ju nämligen att $\dim V(A) + \dim N(A) = 4$ i detta fall, och $\dim V(A) \leq 3$, så $\dim N(A) \geq 1$). En kort räkning visar att $Av = 0 \Leftrightarrow v = t(-1, -2, 1, 2)$, så vi har faktiskt $\dim N(A) = 1$. Dessutom ser vi att $x_1 = 1, x_2 = 2, x_3 = 1$ och $x_4 = 2$. Vi har alltså identifierat reaktionen

$$CH_4 + 2O_2 \to CO_2 + 2H_2O$$

Metoden ovan visar bara hur bevarandet av atomer i en reaktion återspeglar sig i linjär algebra, och nedan är en frivillig övning på detta för den intresserade:

Övning I en process för framställning av ättiksyreanhydrid, $(CH_3CO)_2O$, utgår man ifrån aceton, $(CH_3)_2CO$, och ättiksyra, $C_2H_4O_2$. Som mellan produkter och biprodukter förekommer keten, CH_2CO, metan, kolmonoxid, CO, och vätgas, H_2.

Kan du ge ett förslag på oberoende reaktioner för processen?

Tricket här är att ställa upp en 3×8-matrix vars nollrum ger lösningarna. Kan du förklara följande förslag:

$$\begin{pmatrix} 2 & 4 & 0 & 4 & 2 & 4 & 6 & 6 \\ 0 & 1 & 1 & 2 & 2 & 2 & 3 & 4 \\ 0 & 0 & 1 & 0 & 1 & 2 & 1 & 3 \end{pmatrix}?$$

Nollrummet har dimension $8-3=5$ och en bas ges t.ex. av kolonnerna i matrisen

$$\begin{pmatrix} 2 & 1 & -2 & 1 & -1 \\ -2 & -1 & 0 & -2 & -1 \\ 0 & -1 & -2 & -1 & -3 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Varje rad är en komponent och ordningen mellan komponenterna är densamma som komponenterna i den ursprungliga matrisen. Första kolonnen svarar t.ex. mot reaktionen

$$2CH_4 \to 2H_2 + C_2H_4.$$

Bestäm de övriga.

Ett ekonomi-exempel

Tre hanverkare ska hjälpa till att renovera varandras hus. Alla tre arbetar sammanlagt 10 dagar enigt följande schema:

Snickare: 3 dagar i sitt hus, 4 hos elektrikern och 3 hos rörmokaren
Elektriker: 5 dagar i sitt hus, 2 hos snickaren och 3 hos rörmokaren
Rörmokare: 2 dagar i sitt hus, 4 hos elektrikern och 4 hos snickaren.

Vi kan beskriva detta i följande matris, där kolonnerna representerar hantverkar och raderna vistelse i deras hus:

$$A = \begin{pmatrix} 0.3 & 0.2 & 0.4 \\ 0.4 & 0.5 & 0.4 \\ 0.3 & 0.3 & 0.2 \end{pmatrix}.$$

Här har vi dividerat med 10 för att få elementen att beskriva hur stor del av en dag de tillbringar hos varandra.

Vid prissättning bestämmer hantverkarna att ta en dagslön som gör att var och en tjänar exakt lika mycket på sitt arbete som den måste betala för utfört arbete (notera att alla måste betala till sig själva också). Följande ekvationssystem beskriver då att utgifter är lika med inkomster:

$$\begin{cases} p_S = 0.3p_S + 0.2p_E + 0.4p_R \\ p_E = 0.4p_S + 0.5p_E + 0.4p_R \\ p_R = 0.3p_S + 0.3p_E + 0.2p_R \end{cases}.$$

Det innebär att vektorn $x = (p_S, p_E, p_R)$ ska vara en egenvektor till matrisen A med egenvärdet 1.

Så har då A egenvärdet ett?

$$\begin{vmatrix} \lambda-0.3 & -0.2 & -0.4 \\ -0.4 & \lambda-0.5 & -0.4 \\ -0.3 & -0.3 & \lambda-0.2 \end{vmatrix} = \begin{vmatrix} \lambda-1 & \lambda-1 & \lambda-1 \\ -0.4 & \lambda-0.5 & -0.3 \\ -0.3 & -0.4 & \lambda-0.2 \end{vmatrix} =$$

$$(\lambda-1)\begin{vmatrix} 1 & 1 & 1 \\ -0.4 & \lambda-0.5 & -0.3 \\ -0.3 & -0.4 & \lambda-0.2 \end{vmatrix}.$$

I sista likheten har vi lagt ihop alla kolonner och lagt dem i första. Eftersom alla radsummor är ett följer att vi kan bryta ut $\lambda-1$ ur determinanten, så 1 är ett egenvärde. Detta är alltså alltid sant då radsummorna (eller kolonnsummorna, varför?) är ett!

För att hitta motsvarande egenvektor ska vi lösa ekvationssystemet ovan, vilket vi nu skriver på formen (vi multiplicerar med -10)

$$\begin{cases} -7x_1 + 2x_2 + 4x_3 = 0 \\ 4x_1 - 5x_2 + 4x_3 = 0 \\ 3x_1 + 3x_2 - 8x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = 28t \\ x_2 = 44t \\ x_3 = 27t \end{cases}.$$

(Att vi inte får en entydig lösning beror till dels på att vi inte valt valuta!) Denna lösning talar om hur dagslönen ska förhålla sig för de olika hantverkarna för att ekonomin ska gå ihop!

Anmärkning Detta är ett enkelt exempel på hur nobelpristagaren Wassily Leontief beskrev en ekonomi. Han delar upp den i sektorer (hantverkarna) som handlar med varandra och det vi sett är då att det finns en egenvektor till egenvärdet 1 som definierar vilka priser olika sektorer kan ta ut, för att systemet ska vara i jämvikt.

En viktig observation i vissa praktiska tillämpningar som vi gör från ovan är att

> En matris vars kolonnsummor alla är ett har (minst) en egenvektor med egenvärdet ett

Denna observation återkommer i nästa exempel.

En Markovkedja

En biluthyrningsfirma har tre uthyrningsställen som vi kallar A, B, C. En kund får hyra sin bild på valfritt uthyrningsställe och återlämna bilen på valfritt uthyrningsställe. Efter att ha varit igång i ett år finner man att återlämnande sker på de olika ställena med en sannolikhet som beskrivs i matrisen nedan:

$$P = \begin{pmatrix} 0.8 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.6 \\ 0.1 & 0.5 & 0.2 \end{pmatrix}$$

Här betyder t.ex. elementet 0.6 på plats $(2,3)$ att en bil som hyrs på ställe C med 60% chans återlämnas på ställe B.

Frågan man nu ställer sig är, hur kommer bilarna att fördela sig mellan de olika uthyrningsställena efter en lång tid?

Antag att en bil hyrs på ställe B. Den bilens förväntade historia kan vi då beskriva på följande sätt. Inför vektorn $p_n = (x_n, y_n, z_n)$ som består av sannolikheten x_n att en bil finns på ställe A efter att den lånats ut n gånger, sannolikheten y_n att den är på ställe B och z_n på ställe C. När vi börjar har vi att $p_0 = (0,1,0)$. Sannolikheten att den återlämnas på de olika ställena beskrivs då av $p_1 = Ap_0 = (0.3, 0.2, 0.5)$. Om vi sedan räknar ut $Ap_1 = A^2p_0$ får vi reda på sannolikheten att den hamnar på de olika ställena efter en andra uthyrning. Totalt alltså

$$p_n = A^n p_0.$$

Vi har lärt oss att vi kan beräkna dessa sannolikheter på följande sätt: först bestämmer vi egenvärden (det finns tre olika) och motsvarande egenvektorer, varefter vi skriver

$$A = SDS^{-1}$$

där D är en diagonalmatris med egenvärdena på diagonalen, och S består av egenvektorerna (som är tre olika) tillhörande dessa egenvärden. Sedan kan vi räkna ut $A^n = SD^nS^{-1}$.

Men låt oss räkna så lite som möjligt och istället tänka ut vad som måste gälla. För att förenkla diskussionen räknar vi dock först ut egenvärdena vilka blir

$$\lambda_1 = 1, \lambda_2 = \frac{1}{10}(1+2\sqrt{5}) \approx 0.55, \lambda_3 = \frac{1}{10}(1-2\sqrt{5}) \approx -0.35.$$

Låt egenvektorerna vara S_1, S_2, S_3 till de tre olika egenvärdena $\lambda_1, \lambda_2, \lambda_3$. Då kan vi skriva

$$p_0 = c_1S_1 + c_2S_2 + c_3S_3.$$

Från detta får vi

$$p_1 = Ap_0 = c_1AS_1 + c_2AS_2 + c_3AS_3 = c_1\lambda_1S_1 + c_2\lambda_2S_2 + c_3\lambda_3S_3$$

och fortsätter vi på samma sätt ser vi att

$$p_n = c_1\lambda_1^nS_1 + c_2\lambda_2^nS_2 + c_3\lambda_3^nS_3.$$

Men här gäller att $\lambda_2^n \to 0$ och $\lambda_3^n \to 0$ då $n \to \infty$, och eftersom $\lambda_1 = 1$ så ser vi att

$$p_n \to c_1 S_1.$$

Men eftersom alla p_n har summan ett måste det även gälla $p = c_1 S_1$, så detta är egenvektorn till egenvärdet 1 som är sådan att summan av elementen är ett.

Löser vi det relevanta ekvationssystemet ser vi att $S_1 = t(34, 14, 13)$, och alltså att

$$\lim_{n\to\infty} p_n = \frac{1}{61}(34, 14, 13).$$

Från detta drar vi slutsatsen att bilen efter lång tid kommer att återfinnas på ställe A i $34/61 \approx 56\%$ av fallet, på B i $14/61 \approx 23\%$ av fallen och till C i $13/61 \approx 21\%$ av fallen.

En stunds eftertanke visar att detta faktiskt gäller oavsett var bilen startade. Kan du tänka ut varför?

Här är ett annat exempel att ha som övning:

Övning En entreprenör har just gett sig in i en bransch för att konkurrera med en väletablerad vara. Det visar sig att varje månad så omvänder entreprenörens företag 2 procent av de som använder den etablerade varan. Men det visar sig också att varje månad går 5 procent av entreprenörens kunder tillbaka till den etablerade varan.

Hur lång tid kommer det att ta innan entreprenören har minst 20 procent av marknaden? Vad händer efter lång tid?

Arbeta gärna igenom följande exempel (med dator eller miniräknare) som är en annan illustration av Markovkedjor:

Exempel I England omkring 1950 arbetade man med tre socialgrupper. Data hade visat att sannolikheterna för övergångar (far–son) mellan de olika socialgrupperna kunde beskrivas av följande matris (kolonnerna är fader, raderna son):

$$\begin{pmatrix} 0.448 & 0.054 & 0.011 \\ 0.484 & 0.699 & 0.503 \\ 0.068 & 0.247 & 0.486 \end{pmatrix}.$$

T.ex. gäller att sannolikheten att sonen är i socialgrupp 3 om fadern är socialgrupp 2 lika med 0.247.

Denna matris har kolonnsummor ett och ett egenvärde som är ett med tillhörande egenvektor

$$p = (0.067, 0.624, 0.309).$$

Visa att detta innebär att efter lång tid kommer 62.4% av alla engelsmän att tillhör socialgrupp 2, och bara så lite som 6.7% socialgrupp 1. (Vad är det du ska visa?)

Att lösa andra ordningens linjära differentialekvationer med konstanta koefficienter med hjälp av matriser

Från endim-kursen vet vi att för att lösa differentialekvationen

$$y''(t) + y'(t) - 2y(t) = 0,$$

ska vi först lösa den karakteristiska ekvationen (obs!-samma namn som för egenvärdesekvationen)

$$\lambda^2 + \lambda - 2 = (\lambda - 1)(\lambda + 2) = 0$$

eftersom lösningen då har formen $Ae^{-2t} + Be^t$, varefter vi bestämmer konstanterna. Vi ska nu se hur detta passar in i linjär algebran och det vi gjort.

Tricket är att införa en vektor

$$u(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}$$

och då observera att

$$\begin{pmatrix} y'(t) \\ y''(t) \end{pmatrix} = \begin{pmatrix} y'(t) \\ 2y(t) - y'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}.$$

Detta innebär att

$$u'(t) = Au(t), \quad A = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}.$$

För att lösa detta, antag att vi kan diagonalisera A så att $S^{-1}AS = D$ för någon inverterbar matris S och diagonalmatris D. Om vi då skriver $u(t) = Sw(t)$ för en ny funktion $w(t) = (w_1(t), w_2(t))$, så der vi att

$$w'(t) = S^{-1}u'(t) = S^{-1}Au(t) = S^{-1}ASw(t) = Dw(t).$$

Denna ekvation $w'(t) = Dw(t)$ är bara två oberoende ekvationer

$$w_1'(t) = \lambda_1 w_1(t), \qquad w_2'(t) = \lambda_2 w_2(t),$$

där λ_1, λ_2 är diagnoalelementen i D. Dessa vet vi lösningen på:

$$w_1(t) = Ae^{\lambda_1 t}, \quad w_2(t) = Be^{\lambda_2 t}.$$

Sedan får vi $y(t)$ genom att avläsa första raden i $Sw(t)$.

Vi genomför nu detta. För att se om vi kan diagonalisera ska vi bestämma egenvärdena till A:

$$\det(\lambda I - A) = \begin{vmatrix} \lambda & -1 \\ 2 & \lambda + 1 \end{vmatrix} = \lambda^2 + \lambda - 2 = 0,$$

så egenvärdena är (se ovan) $\lambda = -2, 1$. Motsvarande egenvektorer är (efter en kort räkning) kolonnvektorerna i

$$S = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}$$

och vi har då att $D = S^{-1}AS$ där

$$D = \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Nu har vi siffror till diskussionen ovan och ser då att

$$u(t) = Sw(t) = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} Ae^{-2t} \\ Be^t \end{pmatrix} = \begin{pmatrix} Ae^{-2t} + Be^t \\ -2Ae^{-2t} + Be^t \end{pmatrix}.$$

Men $u(t) = (y(t), y'(t))$, så vi kan avläsa att $y(t) = Ae^{-2t} + Be^t$, precis som vi lärde oss i endim:en.

Exempel Om vi lägger på begynnelsevillkoren $y(0) = 0$, $y'(0) = 1$, hur kan du modifiera räkningarna så att du får den rätta lösningen

$$y(t) = -\frac{1}{3}e^{-2t} + \frac{1}{3}e^t$$

utan att införa konstanter A, B i räkningarna?