

2 Linjär algebra

1949 försökte professorn Wassily Leontief (Harvard) att testa sin matematiska modell för den amerikanska ekonomin. Han använde sig av universitetets mest kraftfulla dator Mark II och matade in data i form av hålkort. Informationen på korten representerade en sammanställning av mer än 250000 olika statistiska data. Leontief hade delat in den amerikanska ekonomin i 500 olika sektorer - kolindustrin, bilindustrin, kommunikationer etc. Varje sektor modellerades med en linjär ekvation som beskrev hur den sektorn samverkade med de övriga 499 sektorerna. Dessvärre kunde inte Mark II hantera det ekvationssystem som blev resultatet, systemet hade 500 ekvationer och lika många variabler vilket var en omöjlig uppgift för denna dator. Leontief tvingades därför att förenkla sin modell till att omfatta 42 ekvationer i lika många variabler. Det tog ändå 56 timmar innan datorn kunde presentera en lösning.

Leontief fick 1973 Nobelpriset i ekonomi för sina matematiska modeller och han var en de första som utforskade tillämpningar av det som vi idag kallar linjär algebra.

I takt med datorutvecklingen har denna matematik blivit allt viktigare och idag resulterar så gott som varje matematisk modell i ett linjärt ekvationssystem i slutändan. Dagens datorer har naturligtvis inte samma begränsningar som 1949 - vi kan nu hantera ekvationssystem med tusentals ekvationer och obekanta. Några exempel på tillämpningsområden följer nedan.

- Den s.k. *minsta kvadratmetoden* är en av de mest använda beräkningsalgoritmerna som existerar. I korthet så går den ut på att finna den "bästa" lösningen till ett överbestämt linjärt ekvationssystem. T.ex. så kan man finna den "bästa" linjära modellen för en datamängd som kommer från en serie experiment.
- Från den datamängd som erhålls vid moderna röntgenmetoder såsom t.ex. magnetröntgen skapas en digital bild av ett tvärsnitt av människokroppen. Denna bild konstrueras via analys av ett överbestämt linjärt ekvationssystem. Minsta kvadratmetoder är viktiga här.
- Molekyler (t.ex. ett protein) kan modelleras genom att bryta ned dem i ett ändligt antal objekt (atomer) som kopplas ihop via attraherande och repellerande krafter. På samma sätt kan t.ex. en bro beskrivas. I båda fallen beskrivs systemets vibrationer av en differentialekvation på matrisform. Matrisens egenvärden svarar mot systemets egenfrekvenser.
- En situation där vissa egenskaper för en art som ärvs från generation till generation kan matematiskt beskrivas med en matrismodell. Liknande modeller som tar hänsyn till olika begränsande och stimulerande faktorer används för studier av olika populationer. En fiskodlare eller en skogsförvaltare kan med hjälp av sådana modeller beräkna optimala nivåer för skörd och sådd.
- I många sammanhang är det intressant att försöka hitta maximum eller minimum av en given funktion $f(x)$ under vissa bivillkor $g(x)=0$. Om f och g är linjära funktioner så resulterar detta i att lösa ett linjärt ekvationssystem. T.ex. kan man försöka hitta den optimala blandningen av komponenter till en produkt, där proportionerna tillåts att variera inom vissa gränser.
- Bild- (t.ex. JPG), video- (t.ex. MPG4) och ljudkomprimeringsalgoritmer (t.ex. MP3) använder Fouriertransformer som är ett exempel på linjära avbildningar. Typiskt så delas materialet upp i små delar som sedan representeras som en vektor. Denna vektor transformeras med en linjär avbildning och man behåller bara den information som är nödvändig för att lura vår hjärna

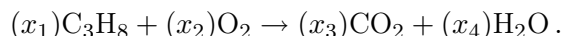
-
- Dynamiska processer med förutbestämda “regler” (som kan bero på slumpen) kan kallas för ett spel. Inom spelteorin försöker man modellera sådana situationer och komma fram till optimala strategier för varje enskild spelare. I dessa matematiska modeller används bl.a. matriser.
 - En datorbild kan representeras som en vektor där varje element svarar mot en pixel. I en flygsimulator eller i ett datorspel där man rör sig i en virtuell 3D-värld (t.ex. det populära Counterstrike) krävs det effektiv implementering av bl.a. rotationer och translationer. Detta utförs med hjälp av matrismultiplikationer.
 - Med metoder från den linjära algebran kan man analysera nätverk av olika slag. Matriser representerar relationer mellan objekten i nätverket. Resultat från matrisalgebran kan sedan överföras på resultat om objekten i nätverket. Som exempel kan nämnas den sociala strukturen i en grupp, planering av flygsträckor och analys av olika mönster i ett riksdagsval.
 - Om man vill konstruera t.ex. en bil ritas man en modell av bilen i en dator. Där interpoleras vissa givna punkter med glatta kurvor (s.k. splines). För att bestämma dessa kurvors matematiska former löser man linjära ekvationssystem.
 - Väderstatistik kan användas för att uppskatta sannolikheten att det snöar en viss dag beroende på om det snöade eller inte dagen innan. Mer generellt så kan man tänka sig ett system som går från ett tillstånd till ett annat med viss sannolikhet. Sådana processer - *Markovprocesser* - hanteras med hjälp av matriser.
 - Kodning används för att bl.a. felkorrigera meddelanden. Kryptering används för att säkert överföra känslig information och har blivit allt viktigare i samband med teknikutvecklingen. I båda fallen omformas data till kodad/krypterad data för att därefter återskapas till originaldata igen. Det finns många olika system som använder sig av matrismetoder för att åstadkomma detta.
 - Teorin för dynamiska system behandlar iterationer av avbildningar eller lösningar till differentialekvationer. Man får en avbildning $T(t)$ vars linjarisering är en matris. Om det största egenvärdet till denna matris växer exponentiellt så har systemet ett kaotiskt uppförande (Exempel på dynamiska system är vårt solsystem, partiklar i en vätska/gas, elektroner i ett plasma etc.)

Nedan presenteras nu några enkla exempel där metoder från den linjära algebran används, i vissa exempel kombinerat med matematisk analys.

2.1 Kemisk balans

Brinnande propangas

Kemiska ekvationer beskriver hur kvantiteter av olika ämnen produceras och konsumeras genom kemiska reaktioner. T.ex. när propangas brinner, reagerar propan (C_3H_8) med syre (O_2) och bildar då koldioxid (CO_2) och vatten (H_2O), enligt en ekvation som har följande form;



För att "balansera" denna ekvation måste en kemist finna positiva heltal x_1, \dots, x_4 sådana att det totala antalet kol-, väte-, och syreatomer på den vänstra sidan motsvaras av ett lika stort antal på den högra sidan (eftersom inga atomer försvinner eller tillkommer genom reaktionen).

Kemisk matematik

En systematisk metod för att balansera sådana kemiska ekvationer är att med hjälp av vektorer ställa upp en ekvation som beskriver antalet atomer av de grundämnen som förekommer i reaktionen. Eftersom den reaktion som beskrivs ovan innehåller tre olika grundämnen - kol, väte och syre - konstruerar vi för varje reagent en vektor i \mathbb{R}^3 som bokför antalet atomer i varje ingående molekyl. Första elementet i vektorn ger antalet kolatomer, andra elementet ger antalet väteatomer och det tredje ger antalet syreatomer enligt följande mönster;

$$\text{C}_3\text{H}_8 : \begin{pmatrix} 3 \\ 8 \\ 0 \end{pmatrix} \quad \text{O}_2 : \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \quad \text{CO}_2 : \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \quad \text{H}_2\text{O} : \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

Insättning av dessa vektorer i ekvationen ovan ger då att heltalen x_1, \dots, x_4 måste uppfylla

$$x_1 \begin{pmatrix} 3 \\ 8 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

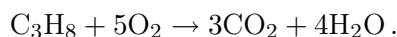
eller

$$x_1 \begin{pmatrix} 3 \\ 8 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} - x_3 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - x_4 \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

vilket är ekvivalent med det linjära ekvationssystemet

$$\begin{cases} 3x_1 - x_3 = 0 \\ 8x_1 - 2x_4 = 0 \\ 2x_2 - 2x_3 - x_4 = 0 \end{cases}$$

Gausselimination ger lösningarna $x_1 = \frac{1}{4}t$, $x_2 = \frac{5}{4}t$, $x_3 = \frac{3}{4}t$, $x_4 = t$ där $t \in \mathbb{R}$ är en parameter. Villkoret att x_1, \dots, x_4 skall vara positiva heltal är uppfyllt om vi t.ex. väljer $t = 4$. Då blir $x_1 = 1$, $x_2 = 5$, $x_3 = 3$ och $x_4 = 4$. Den balanserade kemiska ekvationen blir således



Anmärkning. Den kemiska ekvationen blir balanserad för alla val $t = 4n$ där n är ett positivt heltal men kemister föredrar den ekvation där koefficienterna är så små som möjligt.

2.2 Elektriska kretsar

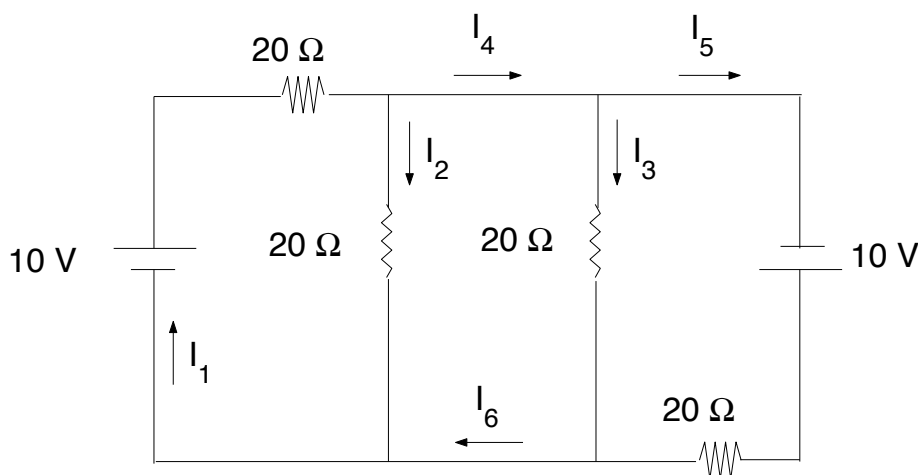
Fysikaliska principer

Här skall vi se vad en elektrisk krets kan ha med linjära ekvationssystem att göra. Den enklaste elektriska kretsen består av spänningskällor och motstånd. Spänningskällorna (t.ex. batterier) alstrar en ström i kretsen och motstånden (t.ex. glödlampor) begränsar strömmens storlek. Strömmens flöde i en krets bestäms av tre grundläggande fysikaliska principer.

1. **Ohms lag.** Om ett motstånd i en krets har resistansen R (mäts i ohm Ω) så gäller sambandet $U = RI$, där U är spänningsfallet över motståndet (mäts i volt V) och I är strömmen (mäts i ampere A).
2. **Kirchoffs första lag.** Summan av ingående strömmar in i en knutpunkt är lika med summan av utgående strömmar i samma knutpunkt.
3. **Kirchoffs andra lag.** Summan av spänningsfallen i en sluten krets är alltid 0.

På med strömmen!

Låt oss tillämpa dessa lagar för att bestämma strömmarna i kretsen nedan.



Tillämpning av Kirchoffs första lag på de fyra knutpunkterna ger ekvationerna

$$\begin{aligned} I_1 &= I_2 + I_4 \\ I_4 &= I_3 + I_5 \\ I_2 + I_6 &= I_1 \\ I_3 + I_5 &= I_6 \end{aligned}$$

Vidare, så ger Kirchoffs andra lag tillsammans med Ohms lag på två slutna kretsar ekvationerna

$$20I_3 - 20I_2 = 0 \text{ och } -10 + 20I_5 - 20I_2 = 0$$

Vi kan nu sammanfatta ekvationerna ovan som det linjära ekvationssystemet

$$\begin{cases} I_1 - I_2 - I_4 = 0 \\ -I_3 + I_4 - I_5 = 0 \\ -I_1 + I_2 + I_6 = 0 \\ I_3 + I_5 - I_6 = 0 \\ 2I_1 + 2I_3 = 1 \\ -2I_2 + 2I_5 = 1 \end{cases}$$

Gausselimination ger sedan $I_1 = I_4 = I_5 = I_6 = 1/2 \text{ A}$, $I_2 = I_3 = 0 \text{ A}$.

2.3 Djuruppfödning

En djuruppfödare är intresserad av att veta hur många djur han kan sälja årligen utan att det påverkar hans totala djurpopulation alltför negativt. Låt oss försöka konstruera en enkel matematisk modell för detta ändamål.

Om vi antar att djuren blir könsmogna efter 1 år och reproducerar sig därefter så är det intressant att veta dels hur många ungdjur (icke könsmogna) som finns och dels hur många vuxna djur (könsmogna) som finns. Låt därför u_k och v_k beteckna antalet ungdjur resp. vuxna djur i djuruppfödarens ägo år k .

Vi antar nu att reproduktionshastigheten beskrivs av en parameter f , dvs. vi antar att från v_k vuxna djur får vi ett tillskott av $f \cdot v_k$ ungdjur året därpå. I samma anda så antar vi vidare att av u_k ungdjur och v_k vuxna djur så avlider $d_u \cdot u_k$ resp. $d_v \cdot v_k$ av dessa under året. Slutligen antar vi att uppfödaren säljer en andel s av de vuxna djuren årligen, dvs. från v_k vuxna djur säljs $s \cdot v_k$.

Sammantaget kan vi nu skriva ned ett samband mellan populationen u_k, v_k år k och populationen u_{k+1}, v_{k+1} år $k+1$ enligt följande.

$$\begin{cases} u_{k+1} = f v_k \\ v_{k+1} = u_k + v_k - d_u u_k - d_v v_k - s u_k \end{cases} \quad \text{eller} \quad \begin{cases} u_{k+1} = f v_k \\ v_{k+1} = (1 - d_u) u_k + (1 - d_v - s) v_k \end{cases}$$

Om vi nu definierar

$$X_k = \begin{pmatrix} u_k \\ v_k \end{pmatrix}, \quad k = 0, 1, 2, \dots, \quad \text{och} \quad A = \begin{pmatrix} 0 & f \\ (1 - d_u) & (1 - d_v - s) \end{pmatrix}$$

så kan ekvationerna formuleras som det eleganta matrissambandet

$$X_{k+1} = A X_k, \quad k = 0, 1, 2, \dots$$

Differensekvationer

Ett matrissamband av typen

$$X_{k+1} = A X_k, \quad k = 0, 1, 2, \dots$$

är ett exempel på ett dynamiskt system som har formen av en s.k. *differensekvation*. En lösning utgörs av den följd av vektorer X_0, X_1, X_2, \dots som är relaterade till varandra enligt differensekvation $X_{k+1} = A X_k$. Dylika differensekvationer uppträder naturligt i många linjära modeller och intressanta frågeställningar är t.ex.

1. För ett givet begynnelsestillstånd X_0 , finns det en vektor X^* sådan att $\lim_{k \rightarrow \infty} X_k = X^*$?
2. Om vi kan visa att gränsvärdet X^* existerar, kan vi då också beräkna X^* ?
3. För ett givet begynnelsestillstånd X_0 , kan vi hitta en formel som låter oss beräkna X_k uttryckt i X_0 ?
4. Givet en vektor Y , kan vi bestämma X_0 så att $X_k = Y$?
5. Givet en vektor X^* , kan vi bestämma X_0 så att $\lim_{k \rightarrow \infty} X_k = X^*$?

En lösningsmetod

Om begynnelsestillståndet X_0 är känt så kan man beräkna X_1, X_2, \dots enligt

$$\begin{aligned} X_1 &= AX_0 \\ X_2 &= AX_1 = A^2 X_0 \\ X_3 &= AX_2 = A^2 X_1 = A^3 X_0 \\ &\dots \\ X_k &= A^k X_0, \quad k = 1, 2, \dots \end{aligned}$$

Alltså har vi en lösningsformel $X_k = A^k X_0$ dvs. vi har i princip besvarat fråga nummer 3 ovan. Problemet med denna formel är att vi måste beräkna (stora) potenser av en matris, A^k och det innebär mycket arbete. De övriga frågorna är också ganska komplicerade att besvara enbart med denna formel. Kan vi möjligen hitta en mer explicit formel? Ja ibland går det, t.ex. om vi antar att matrisen A är diagonaliserbar. Då kan vi skriva

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

där D är den diagonalmatris som har egenvärdena till A som diagonalelement och S är den matris som har motsvarande egenvektorer som kolonner. Fördelen är att nu blir det enkelt att beräkna A^k ;

$$A^k = (SDS^{-1})^k = SDS^{-1}SDS^{-1} \dots SDS^{-1}SDS^{-1} = SD^kS^{-1}$$

dvs. vi behöver bara kunna beräkna stora potenser av D . Men detta är enkelt eftersom D är en diagonalmatris! T.ex. om (i 2×2 -fallet)

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{så är } D^k = \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix}.$$

Vi kan även besvara de två första frågorna nu; gränsvektorn X^* existerar om och endast om $\lim_{k \rightarrow \infty} D^k$ existerar dvs. om och endast om $|\lambda_1| \leq 1$ och $|\lambda_2| \leq 1$.

Åter till farmen

Låt oss nu försöka använda djuruppfödarens modell till att bestämma andelen vuxna djur s som skall säljas på ett sådant sätt att populationen varken dör ut eller växer utan övre gräns. Vi antar att $f = 0.9$, $d_u = 0.1$ samt att $d_v = 0.2$. Då är $X_k = A^k X_0$ där

$$A = \begin{pmatrix} 0 & 0.9 \\ 0.9 & 0.8 - s \end{pmatrix}.$$

Observera att denna matris är symmetrisk och därmed diagonaliserbar. Låt oss kalla egenvärdena till A för λ_1 och λ_2 och samtidigt anta att $|\lambda_1| \geq |\lambda_2|$. Då gäller (enligt diskussionen ovan) att om $|\lambda_1| < 1$ så kommer populationen att dö ut, dvs.

$$X_k \rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Om $|\lambda_1| > 1$ så medför det med ett liknande argument att

$$X_k \rightarrow \begin{pmatrix} \infty \\ \infty \end{pmatrix}$$

dvs. populationen växer obegränsat. Alltså vill vi bestämma andelen sålda djur s så att $|\lambda_1| = 1$. Det karakteristiska polynomet till A ges av

$$p(\lambda) = \det(A - \lambda I) = \lambda^2 - (0.8 - s)\lambda - 0.81.$$

Eftersom $p(1) = s - 0.61$ så kan vi välja $s = 0.61$ och därmed få $\lambda_1 = 1$ och $\lambda_2 = -0.81$. Alltså kan djuruppfödaren sälja 61% av de vuxna djuren utan att populationen riskerar att dö ut eller att population växer okontrollerat. Med denna försäljningsandel kommer populationen att närma sig ett jämnviktsfördelning där det går 9 ungdjur för varje 10 vuxna djur, dvs.

$$\lim_{k \rightarrow \infty} \frac{u_k}{v_k} = \frac{9}{10}.$$

För att visa detta diagonaliserar vi A . Man får $A = SDS^{-1}$ där

$$S = \begin{pmatrix} 9 & 10 \\ 10 & -9 \end{pmatrix} \quad \text{och} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -0.81 \end{pmatrix}.$$

Då kan vi enkelt beräkna

$$A^k = SD^k S^{-1} = \dots = \frac{1}{181} \begin{pmatrix} 81 + 100(-0.81)^k & 90 - 90(-0.81)^k \\ 90 - 90(-0.81)^k & 100 + 81(-0.81)^k \end{pmatrix}$$

så att med

$$X_0 = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$$

får vi

$$\begin{aligned} \lim_{k \rightarrow \infty} X_k &= \lim_{k \rightarrow \infty} A^k X_0 = \lim_{k \rightarrow \infty} SD^k S^{-1} X_0 \\ &= \lim_{k \rightarrow \infty} \frac{1}{181} \begin{pmatrix} 81 + 100(-0.81)^k & 90 - 90(-0.81)^k \\ 90 - 90(-0.81)^k & 100 + 81(-0.81)^k \end{pmatrix} \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \\ &= \frac{1}{181} \begin{pmatrix} 81 & 90 \\ 90 & 100 \end{pmatrix} \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = \frac{1}{181} \begin{pmatrix} 81u_0 + 90v_0 \\ 90u_0 + 100v_0 \end{pmatrix} \end{aligned}$$

Alltså gäller att $X_k \rightarrow X^*$ där

$$X^* = \begin{pmatrix} u^* \\ v^* \end{pmatrix} = \frac{1}{181} \begin{pmatrix} 81u_0 + 90v_0 \\ 90u_0 + 100v_0 \end{pmatrix}$$

och

$$\frac{u^*}{v^*} = \frac{81u_0 + 90v_0}{90u_0 + 100v_0} = \frac{9(9u_0 + 10v_0)}{10(9u_0 + 10v_0)} = \frac{9}{10}$$

vilket bevisar påståendet ovan.

2.4 Google och egenvektorer

Alla inser idag betydelsen av att ha tillgång till en bra sökmotod för att hitta websidor. När vi söker efter en websida vill vi få träffar som är relevanta för vårt sökord. Vad betyder detta? Naturligtvis måste websidan på något sätt innehålla vårt sökord men om det är många sidor som gör det, hur skall vi gradera dem? I ett tidigt skede så räknade sökmotorerna helt enkelt bara hur många gånger det aktuella sökordet förekommer, men denna metod ger ofta mycket dåliga resultat. Vi vill på något sätt kunna avgöra *hur* pass viktig en websida är. En idé är att beräkna hur många *andra* websidor som länkar till den aktuella websidan, men denna metod missar många sidor som kan vara väldigt viktiga utan att för den skull ha ett stort antal länkar till sig. En mer sofistikerad variant av denna idé används av sökmotorn Google (www.google.se). Google mäter graden av viktighet genom att använda ett system som kallas för PageRank och utvecklades av grundarna till Google. På googles hemsida kan man läsa att

PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But, Google looks at more than the sheer volume of votes, or links a page receives; it also analyzes the page that casts the vote. Votes cast by pages that are themselves "important" weigh more heavily and help to make other pages "important."

Lite löst kan man alltså säga att en websida är viktig om andra viktiga websidor har länkar till den. Detta låter som en cirkeldefinition men låt oss försöka se vad denna tanke kan användas till.

Viktade viktigheter

Antag att vi numrerar alla websidor från 1 till n och låter x_k vara ett mått på hur pass viktig websida nummer k är. Tanken med PageRank ovan är då att talet x_k är proportionellt mot summan av alla x_i sådana att sida i har en länk till sida k . Alltså har vi ett system av ekvationer som kan se ut ungefär som

$$\begin{cases} x_1 = \alpha(x_{23} + x_{341} + x_{4525}) \\ x_2 = \alpha(x_3 + x_{634} + x_{346347} + x_{8286352}) \\ \vdots \end{cases}$$

Detta är ett (gigantiskt) linjärt ekvationssystem! Låt nu

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

och definiera $A = (a_{ij})$ som den $n \times n$ -matris som uppfyller att

$$a_{ij} = \begin{cases} 1 & \text{om sida } i \text{ länkar till sida } j \\ 0 & \text{annars} \end{cases}$$

Då kan ekvationssystemet ovan skrivas som

$$X = \alpha AX$$

dvs. ekvationerna beskriver ett egenvärdesproblem! Vektorn X som vi vill ta reda på är tydligen en egenvektor till matrisen A hörande till egenvärdet $\lambda = 1/\alpha$. För att beräkna X kan man

ta reda på samtliga egenvektorer till A och sedan välja X som en egenvektor med enbart positiva element. Om vi har beräknat X så är det sedan lätt att rangordna websidorna i fråga om viktighet; den viktigaste websidan är den sida k som hör ihop med det största talet x_k i vektorn X . På så sätt kan Google ordna sidorna vid en sökning så att den som letar får hjälp med att rangordna sidorna och därmed sparar tid.

Är HV bäst?

Denna idé för att rangordna objekt med hjälp av egenvektorer går tillbaka till 50-talet och har fått allt större betydelse i takt med den ökade graden av webapplikationer. Det finns dock många andra situationer där det kan vara intressant att göra en "intelligent" ranking enligt metoden ovan. Ta t.ex. en turnering där ett antal lag har mött varandra ett antal gånger. Om x_k är styrkan hos lag k och om vi antar att styrkan hos ett lag k är proportionellt mot styrkan i de lag som lag k har besegrat så återfår vi ett ekvationssystem av formen

$$X = \alpha AX$$

där $A = (a_{ij})$ och a_{ij} är antalet gånger som lag i besegrat lag j . Låt oss ta ett litet exempel. Antag att vi har 5 lag som har kämpat i en serie där alla mött varandra två gånger - en hemmamatch och en bortamatch. Matrisen A fick då följande utseende;

$$A = \begin{pmatrix} 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 1 \\ 1 & 1 & 2 & 1 & 0 \end{pmatrix}$$

Här ser vi t.ex. att lag 1 har vunnit två gånger mot lag 2 och 1 gång mot lag 5 osv. Beräkning av egenvektorerna till A ger att

$$X = \begin{pmatrix} 0.38 \\ 0.49 \\ 0.32 \\ 0.50 \\ 0.52 \end{pmatrix}.$$

Alltså är lag 5 det lag som rankas högst och därefter följer i fallande ordning lag 4, lag 2, lag 1 och sist lag 3. Observera att både lag 5 och lag 4 vunnit lika många matcher, men lag 5 har vunnit mot starkare lag jämfört med lag 4. Likaså har lag 1 och lag 3 vunnit lika många matcher men eftersom lag 1 vann en match mot vinnaren lag 5 så rankas lag 1 högre.

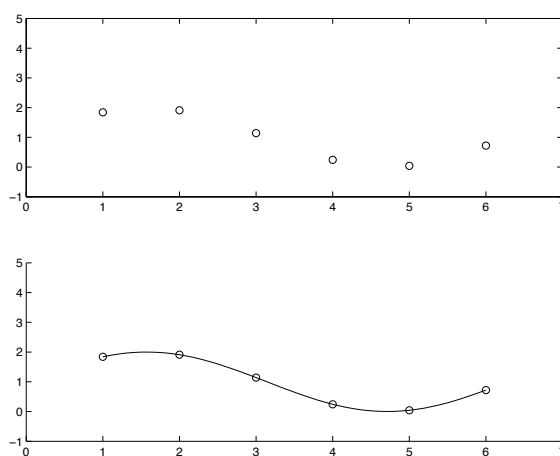
2.5 Designerkurvor

När en designer arbetar med ett CAD-program är en metod att ge som indata ett antal punkter, s.k. *definitions punkter* och sedan låta programmet beräkna och rita ut en kurva som "anpassar sig" till dessa definitionspunkter. Om kurvan passerar genom punkterna så talar vi om en *interpolerande* kurva, och annars om en *approximerande* kurva. Problemet med att finna en interpolerande kurva kan enkelt formuleras som att givet definitionspunkterna

$$(x_1, y_1), \dots, (x_n, y_n)$$

vill vi finna en kurva som passerar genom samtliga n punkter.

Detta är ett klassiskt problem som är lätt att lösa; t.ex. kan man alltid finna ett polynom av grad $n - 1$ som interpolerar dessa punkter. I praktiska sammanhang är detta sällan gångbart eftersom om n är stort så kommer kurvan som bestäms av polynomet att kraftigt oscillera mellan definitionspunkterna. Istället använder man sig av en styckvis definierad kurva dvs. en kurva sammansatt av många delkurvor. Dessa delkurvor är vanligtvis polynom av relativt låg grad och varje delkurva interpolerar bara ett fåtal definitionspunkter.



Kubiska splines

Ordet *spline* är ett engelskt ord för ett rithjälpmiddel som består av en tunn, flexibel träremsa. Den används som en typ av interpolerande linjal genom att fixera den med ett antal stöd i definitionspunkterna som gör det möjligt att rita en kurva som passerar genom dessa punkter. Vi vill nu skapa en matematisk motsvarighet till denna spline. Vi antar för enkelhets skull att interpolationspunkterna är jämnt fördelade i x -led, dvs. vi antar att

$$x_i - x_{i-1} = h, \quad i = 1, \dots, n.$$

Låt $y = S(x)$ vara den interpolerande kurva vi söker, definierad för $x \in [x_1, x_n]$. Vi vill att denna kurva skall beskriva splinelinjalens deformation, när definitionspunkterna är $(x_1, y_1), \dots, (x_n, y_n)$. Om vi ser linjalen som en punktbelastad balk så får vi enligt hållfasthetslärans ekvationer villkoret att fjärde derivatan av $S(x)$ är noll mellan interpolationspunkterna, dvs.

$$S^{(4)}(x) = 0, \quad \text{för } x \in]x_i, x_{i-1}[, \quad i = 1, \dots, n$$

För små böjningar säger också denna hållfasthetslära att $S(x)$ måste ha två kontinuerliga derivator, dvs. $S(x)$, $S'(x)$ och $S''(x)$ måste vara kontinuerliga för $x \in [x_1, x_n]$. Det är det sista villkoret, att $S'''(x)$ är kontinuerlig som gör att vi uppfattar kurvan som "jämn" dvs. vi uppfattar den som en hel kurva utan skarpa kanter någonstans. Matematiskt så säger man att kurvan har en kontinuerlig *krökning*. Detta är ett minikrav men det räcker långt eftersom vi klarar av att se diskontinuiteter i krökningen, dvs. i $S'''(x)$, men diskontinuiteter i högre ordningens derivator märker vi inte med blotta ögat.

Eftersom $S^{(4)}(x) = 0$, i varje delintervall $]x_i, x_{i-1}[$ så finner vi efter fyra integrationer att $S(x)$ måste vara ett tredjegradspolynom på varje delintervall. Härifrån kommer alltså benämningen

kubiska splines för sådana kurvor. Följaktligen kan vi skriva

$$S(x) = \begin{cases} S_1(x), & x_1 \leq x \leq x_2 \\ S_2(x), & x_2 \leq x \leq x_3 \\ \vdots \\ S_{n-1}(x), & x_{n-1} \leq x \leq x_n \end{cases}$$

där $S_1(x), S_2(x), \dots, S_{n-1}(x)$ är kubiska polynom definierade på respektive delintervall.

Beräkning av koefficienterna

Om vi (för att förenkla de kommande beräkningarna) skriver de kubiska polynomen som

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i, \quad i = 1, \dots, n - 1$$

så har vi att bestämma de $4n - 4$ koefficienterna a_i, b_i, c_i, d_i för $i = 1, \dots, n - 1$.

Följande villkor måste vara uppfyllda.

1. $S(x)$ skall interpolera definitionspunkterna (x_i, y_i) , $i = 1, \dots, n$, dvs. $S(x_i) = y_i$, $i = 1, \dots, n$.
2. $S(x)$ är kontinuerlig på $[x_1, x_n]$, dvs. $S_{i-1}(x_i) = S_i(x_i)$, $i = 2, \dots, n - 1$.
3. $S'(x)$ är kontinuerlig på $[x_1, x_n]$, dvs. $S'_{i-1}(x_i) = S'_i(x_i)$, $i = 2, \dots, n - 1$.
4. $S''(x)$ är kontinuerlig på $[x_1, x_n]$, dvs. $S''_{i-1}(x_i) = S''_i(x_i)$, $i = 2, \dots, n - 1$.

Dessa totalt $4n - 6$ villkor ger upphov till lika många ekvationer, alltså saknas det två ekvationer men vi återkommer till detta. Om man genomför beräkningarna så kan resultatet skrivas som

$$\begin{aligned} a_i &= \frac{m_{i+1} - m_i}{6h} \\ b_i &= \frac{m_i}{2} \\ c_i &= \frac{y_{i+1} - y_i}{h} - \frac{(m_{i+1} + 2m_i)h}{6} \\ d_i &= y_i \end{aligned}$$

för $i = 1, \dots, n - 1$. Här är $h = x_i - x_{i-1}$ och vi har uttryckt koefficienterna med hjälp av kvantiteterna $m_i = S''_i(x_i)$, $i = 1, \dots, n$ som entydigt bestämmer våra kubiska splines. Villkoren ovan ger upphov till följande ekvationer

$$\begin{cases} m_1 + 4m_2 + m_3 = 6(y_1 - 2y_2 + y_3)/h^2 \\ m_2 + 4m_3 + m_4 = 6(y_2 - 2y_3 + y_4)/h^2 \\ \vdots \\ m_{n-2} + 4m_{n-1} + m_n = 6(y_{n-2} - 2y_{n-1} + y_n)/h^2 \end{cases}$$

Detta är ett linjärt ekvationssystem med $n - 2$ ekvationer och n obekanta. Precis som vi observerade tidigare så saknar vi två ekvationer för att entydigt bestämma de obekanta m_1, \dots, m_n . Skälet är helt enkelt att det finns oändligt många kubiska splines som uppfyller kraven ovan. För att välja en av dessa möjliga splines måste vi tillföra villkor i ändpunkterna (x_1, y_1) och (x_n, y_n) . Det finns många olika sätt att göra detta på, en vanlig variant är att man kräver att andraderivatorna i ändpunkterna är noll, dvs. vi sätter $m_1 = m_n = 0$. Detta resulterar då i det kvadratiska ekvationssystemet

$$\begin{pmatrix} 4 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 4 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 4 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} m_2 \\ m_3 \\ \vdots \\ m_{n-2} \\ m_{n-1} \end{pmatrix} = \frac{6}{h^2} \begin{pmatrix} y_1 - y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ \vdots \\ y_{n-4} - 2y_{n-3} + y_{n-2} \\ y_{n-3} - 2y_{n-2} + y_{n-1} \\ y_{n-2} - 2y_{n-1} + y_n \end{pmatrix}$$

nu skrivet på matrisform. (Fysikaliskt betyder villkoret $m_1 = m_n = 0$ att den elastiska linjalen fritt får fortsätta i rätta linjer efter ändpunkterna).

Ögats känslighet

Låt oss ta ett exempel. I tabellen nedan har man vid ett experiment uppmätt ögats spektrala känslighet (i enheten lm/W) för ett antal olika våglängder. För vilken våglängd är ögats känslighet som störst?

Känslighet (lm/W)	198	649	587	226	82
Våglängd (nm)	500	540	580	620	660

Alltså är $(x_1, y_1) = (500, 198), \dots, (x_5, y_5) = (660, 82)$ och $h = x_i - x_{i-1} = 40$. Med $m_1 = m_5 = 0$ fås enligt ovan ekvationssystemet

$$\begin{pmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} m_2 \\ m_3 \\ m_4 \end{pmatrix} = \begin{pmatrix} -1.9237 \\ -1.1212 \\ 0.8137 \end{pmatrix}$$

Det följer att $m_2 = -0.4207$, $m_3 = -0.2411$ och $m_4 = 0.2637$ så att vi kan beräkna koefficienterna till polynomen med hjälp av uttrycken ovan. Vår kubiska spline blir

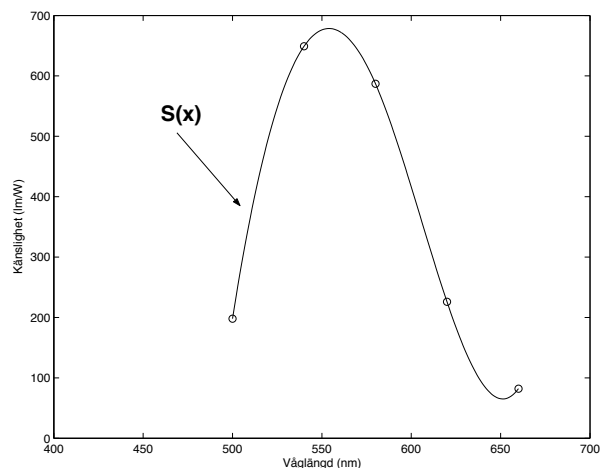
$$S(x) = \begin{cases} -0.0018(x - 500)^3 + 14.08(x - 500) + 198 & , \quad 500 \leq x \leq 540 \\ 0.0007(x - 540)^3 - 0.2103(x - 540)^2 + 5.666(x - 540) + 649 & , \quad 540 \leq x \leq 600 \\ 0.0021(x - 600)^3 - 0.1205(x - 600)^2 - 7.569(x - 600) + 587 & , \quad 600 \leq x \leq 640 \\ -0.0011(x - 640)^3 + 0.1319(x - 640)^2 - 7.116(x - 640) + 226 & , \quad 640 \leq x \leq 680 \end{cases}$$

I figuren till höger plottas mätpunkterna och vår interpolerande kurva. Vi ser att det finns ett största värde i intervallet $[540, 600]$ och för att finna detta deriverar vi $S(x)$ och letar efter stationära punkter. Om $x \in [540, 600]$ så fås

$$S'(x) = 0.0021(x - 540)^2 - 0.4206(x - 540) + 5.666$$

Den rot till ekvationen $S'(x) = 0$ som ligger i det aktuella intervallet visar sig vara $x^* \approx 555$ nm. Ögats maximala ljuskänslighet uppnås alltså för våglängden 550 nm och då är känsligheten

$$S(x^*) \approx 689 \text{ lm/W.}$$



2.6 Temperaturfördelning i en metallskiva

Diskretisering av s.k. kontinuerliga system resulterar ofta i linjära ekvationssystem - vi skall här titta lite närmare på ett sådant exempel. Värmeflödet genom ett 2-dimensionellt objekt, t.ex. en tunn metallskiva, bestäms av värmeledningsekvationen

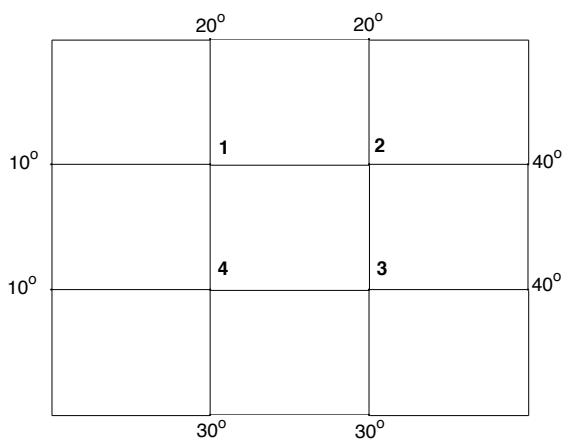
$$\mu \nabla^2 T = \frac{\partial T}{\partial t}.$$

Detta är en s.k. *partiell differentialekvation*; här är temperaturen $T = T(x, y, t)$ en funktion av 3 variabler och $\nabla^2 T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}$. Parametern μ är en värmeledningskonstant som är materialberoende och här låter vi $\mu = 1$. Betrakta nu en kvadratisk tunn metallskiva där vi ser till att de fyra kanternas temperaturer fixeras till 10° , 20° , 30° resp. 40° . Vad blir då temperaturen i en punkt (x, y) inne i skivan? Eftersom systemet strävar mot jämvikt så inträder efter en viss tid en *stationär* temperaturfördelning, dvs. $T(x, y, t) = T(x, y)$ och värmeledningsekvationen övergår i detta jämviktssläge till $\nabla^2 T = 0$. Detta är en extremt viktig partiell differentialekvation som modellerar alla möjliga fysikaliska fenomen inom områden som t.ex. elektromagnetism, hydrodynamik och gravitation. Funktioner som är lösningar till denna ekvation kallas för *harmoniska funktioner*, och de uppfyller en slags medelvärdesprincip - temperaturen i en viss punkt $T(x, y)$ måste vara lika med medelvärdet av temperaturen i omkringliggande punkter. Om detta inte vore fallet så skulle vi ha ett värmeflöde och därmed inte jämvikt. Denna egenskap kan visas matematiskt; om $\nabla^2 T = 0$ så får man om h är ett litet tal att

$$T(x, y) \approx \frac{1}{4} (T(x-h, y) + T(x+h, y) + T(x, y-h) + T(x, y+h)).$$

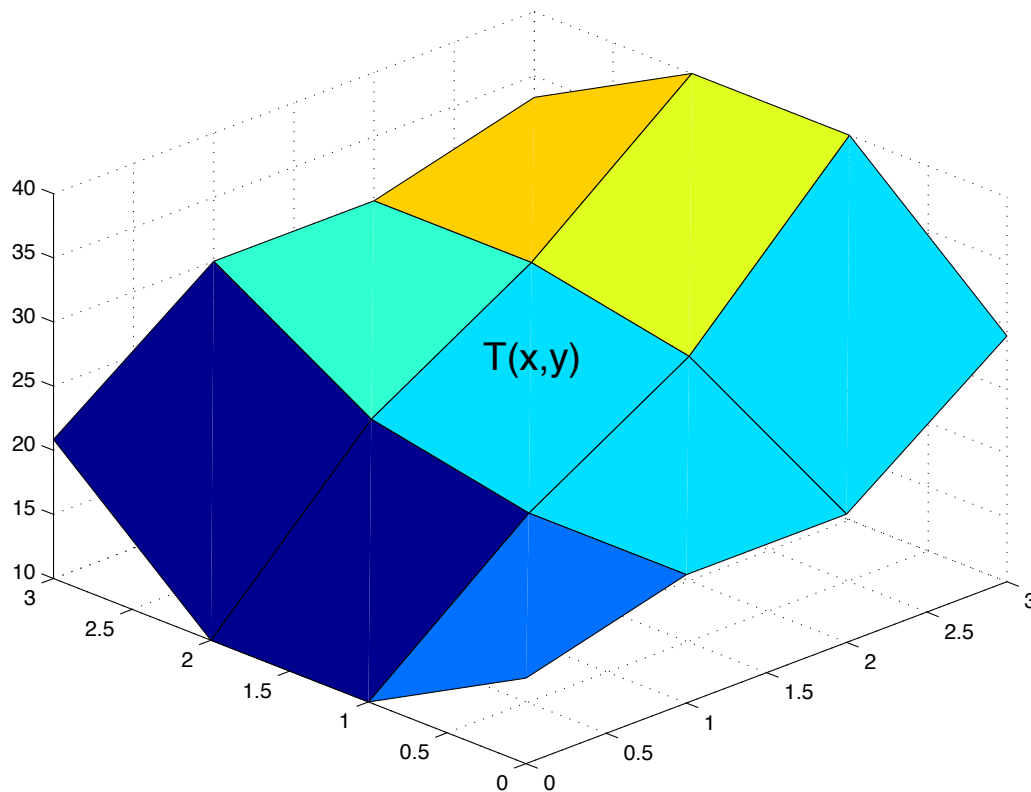
Detta säger bara att $T(x, y)$ är ungefär lika med medelvärdet av temperaturen i de fyra närliggande punkterna $(x-h, y)$, $(x+h, y)$, $(x, y-h)$ och $(x, y+h)$.

Låt oss nu försöka använda linjär algebra för att numeriskt ta fram värmefördelningen i metallskivan. En väldigt enkel modell av situationen visas i figuren; målet är att bestämma temperaturerna i punkterna 1 - 4; kalla dessa temperaturer för T_1, \dots, T_4 . Medelvärdesprincipen ger det linjära ekvationssystemet



$$\begin{cases} T_1 = \frac{1}{4}(10 + 20 + T_2 + T_4) \\ T_2 = \frac{1}{4}(T_1 + 20 + 40 + T_3) \\ T_3 = \frac{1}{4}(T_4 + T_2 + 40 + 30) \\ T_4 = \frac{1}{4}(10 + T_1 + T_3 + 30) \end{cases} \Leftrightarrow \begin{cases} 4T_1 - T_2 - T_4 = 30 \\ -T_1 + 4T_2 - T_3 = 60 \\ -T_2 + 4T_3 - T_4 = 70 \\ -T_1 - T_3 + 4T_4 = 40 \end{cases}$$

Gausselimination ger $T_1 = 20^\circ$, $T_2 = 27.5^\circ$, $T_3 = 30^\circ$, $T_4 = 22.5^\circ$. Nedan visas en plot över temperaturfördelningen i plattan.



Anmärkning. I praktiska sammanhang delar man förstas in området i ett väldigt finmaskigt rutnät - ekvationssystemet kan ha tusentals ekvationer och obekanta. Även för en snabb dator tar det tid att lösa sådana ekvationssystem - man utnyttjar därför mycket sofistikerade metoder för göra beräkningarna så snabbt som möjligt. T.ex. så kan man observera att vårt ekvationssystem ovan kan skrivas på formen $AX = Y$ där

$$X = \begin{pmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{pmatrix}, \quad Y = \begin{pmatrix} 30 \\ 60 \\ 70 \\ 40 \end{pmatrix},$$

och

$$A = \begin{pmatrix} 4 & -1 & 0 & -1 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ -1 & 0 & -1 & 4 \end{pmatrix}.$$

Vi ser att A är en symmetrisk matris med vissa diagonala "band" med nollskilda element, man säger att matrisen är *gles*. Detta kan utnyttjas när man skriver lösningsalgoritmer till ekvationen $AX = Y$.

2.7 Valmatematik

Antag att vi i ett visst land har tre partier att välja mellan när vi röstar - vi representerar partierna med färgerna blå, röd och grön. Resultatet från ett val kan vi då beskriva med en vektor

$$x = \begin{pmatrix} b \\ r \\ g \end{pmatrix}$$

där b , r , g är andelen som röstade på det blå, röda resp. gröna partiet. T.ex. om

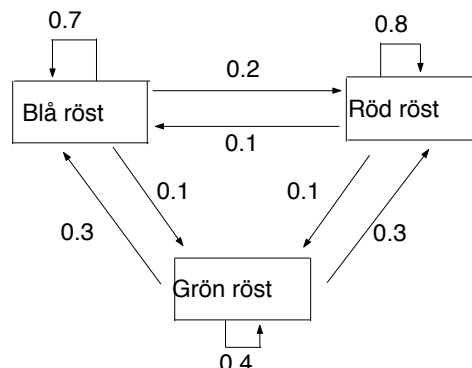
$$x = \begin{pmatrix} 0.35 \\ 0.40 \\ 0.25 \end{pmatrix}$$

så betyder det att 35% röstade blått, 40% röstade rött och slutligen att 25% röstade grönt. Observera att vi här bortser från blankröster och andra mindre partier vilket får konsekvensen att $b + r + g = 1$. En vektor med denna egenskap kallas ibland för *sannolikhetsvektor*.

Markovkedjor

Låt nu x_0, x_1, x_2, \dots vara resultatvektorerna från ett antal efterföljande val. En intressant fråga i detta sammanhang är om det finns något samband mellan resultatet från två intilliggande val, dvs. om vi vet x_k , säger detta någonting om x_{k+1} ?

Med avsikt att försöka besvara denna fråga så analyserar vi statistik från många tidigare val. Då visar det sig att ca. 70% av de som röstade blått i ett visst val också röstade blått i nästkommande val, 20% av dess väljare röstade däremot rött i nästkommande val medans 10% röstade grönt. Vi gör en liknande undersökning för de två andra partierna och resultatet redovisas i figuren här intill.



Med

$$x_k = \begin{pmatrix} b_k \\ r_k \\ g_k \end{pmatrix}, \quad k = 0, 1, 2, \dots$$

kan vi uttrycka dessa samband som linjära ekvationer enligt

$$\begin{cases} b_{k+1} = 0.7b_k + 0.1r_k + 0.3g_k \\ r_{k+1} = 0.2b_k + 0.8r_k + 0.3g_k \\ g_{k+1} = 0.1b_k + 0.1r_k + 0.4g_k \end{cases}$$

eller med matriser $x_{k+1} = Px_k$ där

$$P = \begin{pmatrix} 0.7 & 0.1 & 0.3 \\ 0.2 & 0.8 & 0.3 \\ 0.1 & 0.1 & 0.4 \end{pmatrix}.$$

Observera att P är en matris vars kolonner är sannolikhetsvektorer. En sådan matris kallas för en *stokastisk* matris. Alltså har vi då en följd x_0, x_1, x_2, \dots av sannolikhetsvektorer samt en