4. Egenvärden och invarianta underrum

I detta kapitel utvecklar vi verktyg som hjälper oss att bättre förstå strukturen hos linjära avbildningar.

Som känt så betecknar $L(V)$ mängden av alla operatorer på V, dvs. linjära avbildningar från V till V.

Vi skall nu försöka utreda vad en linjär operator är för något.
Låt alltså $T \in L(V)$, om vi har en direkt summauppdelning

$$V = U_1 \oplus \ldots\ldots \oplus U_m$$

där varje U_j är ett <u>äkta</u> underrum av V, så för att förstå hur T uppför sig behöver vi endast förstå hur varje $T|_{U_j}$ uppför sig på U_j, där $T|_{U_j}$ är restriktionen av T till U_j.

Att jobba med $T|_{U_j}$ borde vara lättare än att jobba med T, ty U_j är ett mindre vektorrum än V.
Men om vi avser att använda verktyg som är nyttiga då man studerar linjära operatorer (som att bilda produkter av operatorer), så stöter vi

på problem:

$T|_{U_j}$ är kanske inte en avbildning från U_j till U_j, dvs. $T|_{U_j} \notin L(U_j)$. Således kan vi endast betrakta uppdelningen $V = U_1 \oplus \ldots \oplus U_m$, där $T(U_j) \subseteq U_j$.

Låt $T \in L(V)$ och U ett underrum av V. Vi säger att U är <u>invariant</u> under T om $u \in U$ implicerar att $Tu \in U$. Alltså U är invariant under T om $T(U) \subseteq U$.

<u>Exempel</u> Låt $T \in L(P_7(\mathbb{R}))$ vara derivata-avbildningen, dvs. $Tp = p'$. Då är $P_4(\mathbb{R})$ ett underrum av $P_7(\mathbb{R})$ och $P_4(\mathbb{R})$ är invariant under T, ty derivatan av ett godtyckligt polynom av gradtal ≤ 4 är också ett polynom av gradtal ≤ 4.

Låt oss nu titta på några enkla exempel av invarianta underrum.

Antag att $T \in L(V)$. Klart att $\{0\}$ är invariant under T. Vidare är det klart att också V är invariant under T.

<u>Fråga</u>: Har varje $T \in L(V)$ alltid något invariant underrum annat än $\{0\}$ och V?

Vi skall senare se att denna fråga har ett <u>jakande</u> svar för operatorer på <u>komplexa</u> vektorrum med dimensionen större än 1 och också för operatorer på <u>reella</u> vektorrum med dimensionen större än 2.

Om $T \in L(V)$, så gäller att

$$T(N(T)) \subseteq N(T)$$

och

$$T(R(T)) \subseteq R(T).$$

Fastän $N(T)$ och $R(T)$ är invarianta underrum för T, så behöver de inte nödvändigtvis förse oss med kätta svar på vår fråga, huruvida det existerar <u>andra</u> invarianta underrum än $\{0\}$ och V. Speciellt gäller ju att om T är inverterbar, så är $N(T) = \{0\}$ och $R(T) = V$.

Vi börjar med att undersöka de enklaste möjliga icke-triviala invarianta underrummen, d.v.s. invarianta underrum med dimensionen 1.

Underrum av V med dimensionen 1 är lätta att beskriva. Tag godt. $u\in V$ med $u\neq 0$ och låt

$$U=\{\alpha u:\alpha\in\mathbb{K}\}.$$

Då är U ett underrum av V med $\dim U=1$ och varje 1-dimensionellt underrum är av denna form.

Om $0\neq u\in V$ och $U=\{\alpha u:\alpha\in\mathbb{K}\}$ är invariant under $T\in L(V)$, så gäller att $Tu\in U$, dvs. det existerar ett $\lambda\in\mathbb{K}$ så att $Tu=\lambda u$.
Omvänt, om $0\neq u\in V$ så att $Tu=\lambda u$ för ngt $\lambda\in\mathbb{K}$, så är $U=\{\alpha u:\alpha\in\mathbb{K}\}$ ett 1-dimensionellt underrum av V som är invariant under T.

Ekvationen

$$Tu=\lambda u$$

är således nära sammanhängande med 1-dimensionella invarianta underrum, och den är så viktig att vi ger ett speciellt namn åt vektorn u och skalären λ som satisfierar den.

En skalär $\lambda \in K$ kallas ett <u>egenvärde</u> till $T \in L(V)$ om det finns en vektor $u \in V$, $u \neq 0$, så att $Tu = \lambda u$.

Nu gäller

$$Tu = \lambda u \Leftrightarrow (T - \lambda I)u = 0,$$

så λ är ett egenvärde till T om och endast om $T - \lambda I$ <u>inte</u> är injektiv.

Alltså T har ett 1-dimensionellt invariant underrum $\Leftrightarrow$ T har ett egenvärde.

Antag $T \in L(V)$ och $\lambda \in K$ är ett egenvärde till T.
En vektor $u \in V$ kallas en <u>egenvektor</u> till T om $Tu = \lambda u$.

Emedan $Tu = \lambda u \Leftrightarrow (T - \lambda I)u = 0$, så följer att

$$N(T - \lambda I) = \{\text{alla egenvektorer till } T \text{ svarande mot } \lambda\},$$

som är ett underrum av V.

Exempel Betrakta $T \in L(\mathbb{K}^2)$ definierad genom

$$T(w,z) = (-z, w).$$

Om $\mathbb{K} = \mathbb{R}$, så har denna avbildning en geometrisk tolkning: T är en moturs rotation med $90°$ kring origo i $\mathbb{R}^2$.
Vi söker $\lambda \in \mathbb{K}$ med

$$(-z, w) = \lambda (w,z) \quad \text{och} \quad (0,0) \neq (w,z) \in \mathbb{K}^2.$$

Alltså

$$-z = \lambda w \quad \text{och} \quad w = \lambda z,$$

varför $-z = \lambda^2 z$. Då $z \neq 0$, får att $\lambda^2 = -1$.
Då $\mathbb{K} = \mathbb{R}$ saknar T egenvärden. Men om $\mathbb{K} = \mathbb{C}$ så är $\lambda = i$ eller $\lambda = -i$. Egenvektorn som svarar mot i är av formen $(a, -ia)$, $a \in \mathbb{C}$, och egenvektorn som svarar mot egenvärdet $-i$ är vektorn av formen (a, ia) med $a \in \mathbb{C}$.

Vi skall nu visa att egenvektorer olika noll svarande mot olika egenvärden är linjärt oberoende.

Sats 4.1 Låt $T \in L(V)$. Antag att $\lambda_1, \dots, \lambda_n$ är olika egenvärden till T och att $v_1, \dots, v_n$ är de motsvarande egenvektorerna som alla är olika noll. Då är $v_1, \dots, v_n$ linjärt oberoende.

Bevis: Antag att $u_1, \dots, u_n$ är linjärt beroende. Låt k vara det minsta positiva heltal så att $u_k \in [u_1, \dots, u_{k-1}]$, dvs. $u_1, \dots, u_{k-1}$ är linjärt oberoende. Då finns entydigt bestämda skalärer $c_1, \dots, c_{k-1}$ så att

$$u_k = c_1 u_1 + \dots + c_{k-1} u_{k-1}.$$

Nu fås

$$\lambda_k u_k = T u_k = c_1 \lambda_1 u_1 + \dots + \lambda_{k-1} c_{k-1} u_{k-1},$$

varför

$$0 = c_1(\lambda_k - \lambda_1) u_1 + c_2(\lambda_k - \lambda_2) u_2 + \dots + c_{k-1}(\lambda_k - \lambda_{k-1}) u_{k-1}.$$

Emedan $\lambda_k \neq \lambda_i$, $i = 1, \dots, k-1$ och $u_1, \dots, u_{k-1}$ är linjärt oberoende, följer att

$$c_1 = c_2 = \dots = c_{k-1} = 0.$$

Detta betyder att $u_k = 0$, vilket strider mot antagandet att alla $u_1, u_2, \dots, u_n$ är olika noll.

Korollarium 4.2 Varje operator på V har högst $\dim V$ olika egenvärden.

Bevis: Låt $T \in L(V)$ och antag att $\lambda_1, \dots, \lambda_n$ är olika egenvärden till T. Låt $u_1, \dots, u_n$ vara motsvarande egenvektorer, som alla är olika noll. Sats 4.1 ger att $u_1, \dots, u_n$ är linjärt oberoende. Alltså $n \leq \dim V$, som önskades.

<u>Exempel</u> Antag att A är en $n\times n$ matris med elementen $a_{ij}\in K$. ~~Ett~~ Ett tal $\lambda \in K$ kallas ett egenvärde till A om det finns en $n \times 1$ matris (kolonnvektor) $X \neq 0$ sådan att

$$AX = \lambda X.$$

~~A~~ Alla sådana $n\times 1$ matriser (kolonnvektorer) $X \neq 0$ kallas för egenvektorer till A svarande mot egenvärdet λ.

Om $\lambda \in K$ är ett egenvärde till A, så gäller

$$(A-\lambda I)X = 0,$$

där I är $n\times n$ enhetsmatrisen. Med andra ord är λ ett egenvärde till A om och endast om det homogena ekvationssystemet $(A-\lambda I)X=0$ har icke-triviala lösningar $X \neq 0$, vilket inträffar då och endast då

$$p(\lambda) = \det(A-\lambda I) = 0,$$

där $\det(A-\lambda I)$ betecknar determinanten av matrisen $A-\lambda I$ och är ett polynom $p(\lambda)$ av gradtal n. Polynomet $p(\lambda)$, med koefficienter i K, kallas karakteristiska polynomet till A.
Nollställena i K för $p(\lambda)$ utgör mängden av egenvärden till A i K.

Om exempelvis $A = \begin{pmatrix} 7 & 8 \\ 1 & 5 \end{pmatrix}$ har vi det karakteristiska polynomet

$$p(\lambda) = \det(A-\lambda I) = \begin{vmatrix} 7-\lambda & 8 \\ 1 & 5-\lambda \end{vmatrix} = (7-\lambda)(5-\lambda) - 8$$
$$= \lambda^2 - 12\lambda + 27 = (\lambda-3)(\lambda-9),$$

med nollställena $\lambda_1 = 3$ och $\lambda_2 = 9$, som utgör egenvärdena till A i $\mathbb{R}$ (och $\mathbb{C}$).

För $\lambda = 3$ är $(2,-1)^T$ en egenvektor, och för $\lambda = 9$ är $(4,1)^T$ en egenvektor.

Om $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, så har A inga reella egenvärden, ty $p(\lambda) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$, varför $\lambda = \pm i$ är egenvärdena till A i $\mathbb{C}$.

<u>Sats 4.3</u> Antag att $T \in L(V)$ och att $M(T)$ är matrisen för T med avseende på <u>någon</u> bas för V. Då har T och $M(T)$ samma egenvärden.

<u>Bevis</u>: Låt $u_1, \ldots, u_n$ vara basen i V för vilken $M(T)$ är matrisframställningen av T. Låt $\lambda \in \mathbb{K}$. Vi visar att λ är ett egenvärde till T om och endast om λ är ett egenvärde till $M(T)$.

Antag att $u \in V$, $u \neq 0$, är sådan att $Tu = \lambda u$. Vi har framställningarna:

$$u = c_1 u_1 + \ldots + c_n u_n, \quad c_1, \ldots, c_n \in \mathbb{K}$$

$$T u_k = a_{1,k} u_1 + \ldots + a_{n,k} u_n, \quad k = 1, \ldots, n$$

$$M(T) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}, \quad a_{ij} \in K$$

$$M(u) = (c_1, \ldots, c_n)^T \neq (0, \ldots, 0).$$

Enligt Sats 2.6 gäller:

$$M(T)M(u) = M(Tu) \overset{Tu=\lambda u}{=} M(\lambda u) = \lambda M(u),$$

så λ är ett egenvärde till $M(T)$.

Omvänt, antag att λ är ett egenvärde till $M(T)$. Låt $X \neq 0$ vara en $n \times 1$ matris med $M(T)X = \lambda X$. Låt $X = (c_1, \ldots, c_n)^T$ och sätt

$$u = c_1 u_1 + \ldots + c_n u_n \in V, \ u \neq 0.$$

Då gäller

$$Tu = \sum_{j=1}^{n} c_j T u_j = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{n} a_{ij} u_i \right)$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} c_j \right) u_i \underset{M(T)X=\lambda X}{=} \sum_{i=1}^{n} (\lambda c_i) u_i = \lambda u,$$

så λ är ett egenvärde till T.

Uppåt triangulära matriser och operatorpolynom

Vi skall nu fortsätta att undersöka operatorer. Att studera operatorer utgör den djupaste och viktigaste delen av linjära algebran. Många av nyckelresultaten inom linjära algebran gäller *inte* för oändligtdimensionella vektorrum, så i fortsättningen antar vi att alla rum som vi arbetar med är ändligtdimensionella.

Låt $T \in L(V)$. Då är också TT väldefinierad och $TT \in L(V)$. Vi skriver ofta T^2 istället för TT. Mera allmänt, om m är ett positivt heltal, så definieras T^m genom

$$T^m := TT \cdots T \quad (m \text{ st}).$$

Vi definierar $T^0 = I \in L(V)$. Om T är inverterbar, så betecknas som känt inversen till T med T^{-1}. Om m är ett positivt heltal, så definierar vi

$$T^{-m} := (T^{-1})^m = T^{-1}T^{-1} \cdots T^{-1} \quad (m \text{ st})$$

Om $T \in L(V)$, så gäller

$$T^m T^n = T^{m+n}, \quad (T^n)^m = T^{nm},$$

där m och n tillåts vara godtyckliga heltal om T är inverterbar och icke-negativa heltal om T <u>inte</u> är inverterbar.

Om $T \in L(V)$ och $p \in P(\mathbb{K})$ är ett givet polynom genom

$$p(t) = c_0 + c_1 t + \dots + c_m t^m$$

för $t \in \mathbb{K}$, då är $p(T)$ en operator definierad genom

$$p(T) = c_0 I + c_1 T + \dots + c_m T^m.$$

<u>Ex.</u> Om $p(t) = t^2$, $t \in \mathbb{K}$, så är $p(T) = T^2$.

För givet $T \in L(V)$ gäller att avbildningen

$$\begin{aligned} P(\mathbb{K}) &\longrightarrow L(V) \\ p &\longmapsto p(T) \end{aligned}$$

är linjär (hemuppgift). Om $p, q \in P(\mathbb{K})$, då är $pq \in P(\mathbb{K})$ definierad genom

$$(pq)(t) = p(t)\,q(t), \quad t \in \mathbb{K}.$$

Om $T \in L(V)$ så gäller $(pq)(T) = p(T)\,q(T)$ för alla $p, q \in P(\mathbb{K})$ (hemuppgift). Notera att två polynom i T kommuterar i betydelsen att $p(T)q(T) = q(T)p(T)$, emedan

$$p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T).$$

Nu kommer vi till ett av de mest centrala resultat rörande operatorer $T \in L(V)$ på komplexa vektorrum.

Sats 4.4 Varje operator $T \in L(V)$ på ett komplext vektorrum $V \neq \{0\}$ har ett egenvärde.

Bevis: Antag att $\dim V = n > 0$ och låt $T \in L(V)$. Välj en vektor $x \neq 0$ i V. Då är

$$x, Tx, \ldots, T^n x,$$

inte linjärt oberoende, ty $\dim V = n$ och vi har $n+1$ vektorer. Således existerar $c_1, \ldots, c_n \in \mathbb{C}$ inte alla noll, och $c_0 \in \mathbb{C}$ med

$$c_0 x + c_1 Tx + \ldots + c_n T^n x = 0.$$

Vi tar de komplexa talen $c_0, \ldots, c_n$ och bildar det komplexa polynomet, som inte är konstant,

$$c_0 + c_1 t + \ldots + c_n t^n, \quad t \in \mathbb{C}.$$

Då får vi enligt Korollarium 3.5, att

$$c_0 + c_1 t + \ldots + c_n t^n = c(t-\lambda_1) \cdots (t-\lambda_m),$$

där $c \neq 0$, $c \in \mathbb{C}$, $m \geq 1$ och varje $\lambda_j \in \mathbb{C}$ och ekvationen gäller för varje $t \in \mathbb{C}$. Nu gäller

$$0 = (c_0 I + c_1 T + \dots + c_n T^n) x$$

$$= c(T - \lambda_1 I)(T - \lambda_2 I) \dots (T - \lambda_m I) x,$$

vilket betyder att för åtminstone ett j är $T - \lambda_j I$ inte injektiv, dvs. T har ett egenvärde.

Betrakta den kvadratiska matrisen

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}$$

vars diagonal består av elementen $a_{1,1}, \dots, a_{n,n}$. En matris kallas uppåt triangulär om alla elementen nedanför diagonalen är 0.

Ex. Matrisen $\begin{pmatrix} 4 & 2 & -7 \\ 0 & 8 & 2 \\ 0 & 0 & 5 \end{pmatrix}$ är en uppåt triangulär matris.

Vanligen kommer vi att återge en uppåt triangulär matris på formen
$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

där 0 i matrisen betyder att alla elementen nedanför diagonalen i denna $n \times n$ matris är 0. Vidare betecknar $*$ att vi inte vet vilka elementen ovanför diagonalen är eller att det är irrelevant vilka elementen är för de problem vi är intresserade av.

Ett centralt mål i linjära algebran är att visa, givet en operator $T \in \mathcal{L}(V)$, att det finns en bas i V med avseende på vilken $M(T)$ är förnämligen enkel.

För att göra denna vaga formulering något mera konkret, strävar vi att åstadkomma att många av elementen i matrisen $M(T)$ är 0.

För stort n, gäller att en uppåt triangulär $n \times n$ matris har nästan hälften av alla sina element lika med 0.

Följande sats visar ett nyttigt samband mellan uppåt triangulära matriser och invarianta underrum.

<u>Sats 4.5</u> Låt $T \in \mathcal{L}(V)$ och låt $v_1, \ldots, v_n$ vara en bas i V. Följande påståenden är ekvivalenta:

(a) Matrisen $M(T)$ av T med avseende på $v_1, \ldots, v_n$ är uppåt triangulär.

(b) $T v_j \in [v_1, \ldots, v_j]$ för varje $j = 1, \ldots, n$.

(c) $[v_1, \ldots, v_j]$ är invariant under T för varje $j = 1, \ldots, n$.

<u>Bevis:</u> (a) $\Leftrightarrow$ (b) följer lätt från definitionerna och lite eftertanke.

(c) $\Rightarrow$ (b) är klart.

Återstår att visa att (b) $\Rightarrow$ (c). Antag att (b) gäller. Fixera $j \in \{1, \ldots, n\}$. Från (b) följer att

$$T v_1 \in [v_1] \subseteq [v_1, \ldots, v_j],$$
$$T v_2 \in [v_1, v_2] \subseteq [v_1, \ldots, v_j],$$
$$\vdots$$
$$T v_j \in [v_1, \ldots, v_j] \subseteq [v_1, \ldots, v_j].$$

Tag godt. $v \in [v_1, \ldots, v_j]$, dvs. $v = \sum_{i=1}^{j} c_i v_i$, där $c_i \in K$. Då

$$T v = \sum_{i=1}^{j} c_i T v_i \in [v_1, \ldots, v_j],$$

dvs. $[v_1, \ldots, v_j]$ är invariant under T, och satsen är bevisad.

<u>Fråga</u>: Kan man genom att titta på matrisframställningen $M(T)$ av en operator $T \in L(V)$ avgöra huruvida T är inverterbar?

Om vi har tur och har en bas med avseende på vilken T har en uppåt triangulär matrisframställning $M(T)$, så blir detta problem lätt, som följande sats visar.

Sats 4.6 Antag att $T \in L(V)$ har en uppåt triangulär matris med avseende på någon bas i V. Då är T inverterbar om och endast om diagonalelementen i den uppåt triangulära matrisen alla är skilda från noll.

Bevis: Antag att $v_1, \dots, v_n$ är en bas i V med avseende på vilken matrisframställningen av T ges av

$$(*) \qquad M(T, v_1, \dots, v_n) = \begin{pmatrix} \lambda_1 & & * \\ & \lambda_2 & \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Vi bör visa att T inte är inverterbar om och endast om något $\lambda_j = 0$.

Antag först att ett $\lambda_j = 0$. Om $\lambda_1 = 0$, så är $Tv_1 = 0$ enligt $(*)$. Men då är T inte inverterbar.

Antag att $1 < j \leq n$ och $\lambda_j = 0$. Från $(*)$ ser vi att

$$\begin{aligned} Tv_1 &\in [v_1, \dots, v_{j-1}] \\ &\vdots \\ Tv_{j-1} &\in [v_1, \dots, v_{j-1}] \end{aligned}$$

samt att $Tv_j \in [v_1, \dots, v_{j-1}]$ då $\lambda_j = 0$. Således kan vi definiera en linjär avbildning

$$S : [v_1, \dots, v_j] \longrightarrow [v_1, \dots, v_{j-1}]$$

genom $Sv = Tv$ för $v \in [v_1, \dots, v_j]$. Antag

S är T:s restriktion till $[v_1, \dots, v_j]$.
Notera att $\dim[v_1, \dots, v_j] = j$ och $\dim[v_1, \dots, v_{j-1}] = j-1$.
Korollarium 2.4 ger att det inte finns någon injektiv linjär avbildning från $[v_1, \dots, v_j]$ till $[v_1, \dots, v_{j-1}]$.
Således existerar ett $0 \neq v \in [v_1, \dots, v_j]$ med $Sv = 0$. Alltså $Tv = 0$, varför T inte är inverterbar.

Omvänt, antag att T inte är inverterbar.
Enligt Sats 2.11 är T inte injektiv. Således existerar $0 \neq v \in V$ med $Tv = 0$.
Då kan v skrivas i formen

$$v = c_1 v_1 + \dots + c_k v_k, \quad \text{där } c_1, \dots, c_k \in K$$

och $c_k \neq 0$ genom att välja k att vara det största indexet med $c_k \neq 0$. Detta är möjligt då $v \neq 0$. På grund av $(*)$ kan vi skriva

$$Tv_k = w + \lambda_k v_k \quad \text{för ngt } w \in [v_1, \dots, v_{k-1}].$$

Alltså

$$\begin{aligned} 0 = Tv &= T(c_1 v_1 + \dots + c_k v_k) \\ &= \underbrace{c_1 T v_1 + \dots + c_{k-1} T v_{k-1}}_{:= u} + c_k T v_k \\ &= u + c_k w + c_k \lambda_k v_k, \end{aligned}$$

där $u \in [v_1, \dots, v_{k-1}]$ enligt Sats 4.5. Emedan

både $v, w \in [v_1, \ldots, v_{k-1}]$ ger ovanstående ekvation (7)
att

$$c_k \lambda_k v_k \in [v_1, \ldots, v_{k-1}].$$

Härur följer att $c_k \lambda_k = 0$ emedan $v_1, \ldots, v_n$ är linjärt oberoende. Då $c_k \neq 0$ fås att $\lambda_k = 0$, och satsen är bevisad.

Tyvärr finns ingen metod att exakt beräkna egenvärdena till en typisk operator från dess matrisframställning med avseende på en godtycklig bas. Det existerar kraftfulla numeriska tillvägagångssätt för att hitta goda approximationer av egenvärdena till en operator från dess matrisframställning.
Men om vi har tur att hitta en bas med avseende på vilken matrisframställningen av operatorn är en uppåt triangulär matris, så blir problemet att beräkna egenvärdena trivialt som följande sats visar.

Sats 4.7 Antag att $T \in L(V)$ har en uppåt triangulär matris med avseende på **någon** bas i V. Då består egenvärdena till T exakt av diagonalelementen i denna uppåt triangulära matris.

Bevis: Antag att $v_1, \ldots, v_n$ är en bas i V med avseende på vilken T har en uppåt triangulär matrisframställning

$$M(T, v_1, \ldots, v_n) = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Låt $\lambda \in K$. Då är

$$M(T-\lambda I, v_1, \ldots, v_n) = \begin{pmatrix} \lambda_1-\lambda & & & * \\ & \lambda_2-\lambda & & \\ & & \ddots & \\ 0 & & & \lambda_n-\lambda \end{pmatrix}.$$

Således är $T-\lambda I$ ej inverterbar om och endast om $\lambda = \lambda_j$ för något j enligt Sats 4.6, dvs. λ är ett egenvärde till T om och endast om $\lambda = \lambda_j$ för något j.

Antag att U och W är underrum av V med

$$V = U \oplus W.$$

Varje $v \in V$ kan <u>entydigt</u> skrivas som

$$v = u + w,$$

där $u \in U$ och $w \in W$. Med denna framställning definierar vi $P_{U,W}$ genom $P_{U,W}(v) = u$. Ofta kallar man $P_{U,W}$ <u>projektionen</u> på U med nollrummet W.

Då är $P_{U,W}$ en linjär avbildning och $P_{U,W}(V) = U$. Alltså $P_{U,W} \in L(V,U)$ eller $P_{U,W} \in L(V)$, då $U \subseteq V$. Om vi byter om U:s och W:s roller i ovanstående framställning, så får vi $P_{W,U}(v) = w$. Således är

$$v = P_{U,W}(v) + P_{W,U}(v) \text{ för alla } v \in V.$$

Notera att om $u_1, \ldots, u_n$ är en bas i U och $w_1, \ldots, w_m$ är en bas i W, så är

$$P_{U,W}(\alpha_1 u_1 + \ldots + \alpha_n u_n + \beta_1 w_1 + \ldots + \beta_m w_m) = \alpha_1 u_1 + \ldots + \alpha_n u_n$$

för alla $\alpha_i, \beta_j \in \mathbb{K}$. Vidare gäller att $P_{U,W}^2 = P_{U,W}$ och att $N(P_{U,W}) = W$.

Sats 4.8 Antag att V är ett komplext vektorrum och att $T \in L(V)$. Då existerar det en bas i V med avseende på vilken T har en uppåt triangulär matrisframställning.

Bevis: Vi visar denna sats genom induktion över $\dim V$.
Om $\dim V = 1$, så är det klart att resultatet gäller.
Antag att $\dim V > 1$ och att det önskade resultatet gäller för alla komplexa vektorrum med dimensionen mindre än $\dim V$.

Låt λ vara ett egenvärde till T. Sats 4.4 garanterar att ett sådant existerar. Låt $0 \neq v \in V$ med $Tv = \lambda v$. Sätt

$$U = \{\alpha v : \alpha \in \mathbb{C}\}.$$

Då är $\dim U = 1$. Enligt Sats 1.14 existerar ett underrum W av V med

$$V = U \oplus W.$$

Emedan $\dim W < \dim V$, skall vi använda vårt induktionsantagande på $T|W$.
Men W behöver inte vara invariant under T, vilket betyder att $T|W$ inte är en operator från W till W. Därför komponerar vi $T|W$ med projektionen $P_{W,U}$ på W, dvs. vi definierar $S \in L(W)$ genom

$$Sw = P_{W,U}T(w) \text{ för } w \in W.$$

Enligt vårt induktionsantagande existerar det en bas $w_1, \ldots, w_{n-1}$ i W sådan ~~att~~ S har en uppåt triangulär matrisframställning med avseende på denna bas.

Vi slutför beviset genom att visa att T har en uppåt triangulär matrisframställning med avseende på basen $v, w_1, \ldots, w_{n-1}$ i V.

Emedan v är en egenvektor till T gäller det att $Tv \in [v]$. Antag nu att $j \in \{1, \ldots, n-1\}$. Då gäller

$$Tw_j = p_{U,W}(Tw_j) + p_{W,U}(Tw_j)$$

$$= p_{U,W}(Tw_j) + Sw_j.$$

Klart att $p_{U,W}(Tw_j) \in U$, varför $p_{U,W}(Tw_j) = \mu v$ för något $\mu \in \mathbb{C}$.

Sats 4,5 ger att $Sw_j \in [w_1, \ldots, w_j]$. Alltså

$$Tw_j \in [v, w_1, \ldots, w_j].$$

Emedan $Tv \in [v]$ och $Tw_j \in [v, w_1, \ldots, w_j]$ ger Sats 4.5 att T har en uppåt triangulär matris med avseende på basen $v, w_1, \ldots, w_{n-1}$, och satsen är bevisad.

Invarianta underrum på reella vektorrum

Vi vet att varje operator på ett komplext vektorrum har ett egenvärde (Sats 4.4). Operatorn $T \in L(\mathbb{R}^2)$ definierad genom $T(x,y) = (-y,x)$ saknar egenvärden, så detta resultat gäller inte för reella vektorrum. Med andra ord, en operator på ett reellt vektorrum olika $\{0\}$ kan sakna invarianta underrum med dimensionen 1.

Emellertid gäller följande resultat.

Sats 4.9 Varje operator från ett ändligtdimensionellt reellt vektorrum olika $\{0\}$ till sig själv har ett invariant underrum med dimensionen 1 eller 2.

Bevis: Antag att V är ett reellt vektorrum med $\dim V = n > 0$ och $T \in L(V)$.
Tag en vektor $v \neq 0$ i V. Då är

$$v, Tv, T^2v, \ldots\ldots, T^nv,$$

inte linjärt oberoende, ty vi har $n+1$ vektorer och $\dim V = n$. Således existerar reella tal $c_0, \ldots, c_n$, alla inte noll, så att

$$c_0v + c_1Tv + \ldots\ldots + c_nT^nv = 0.$$

Vi bildar det reella polynomet

$$c_0 + c_1 t + \dots + c_n t^n, \; t \in \mathbb{R},$$

och får att

$$c_0 + c_1 t + \dots + c_n t^n = c(t-\lambda_1) \dots (t-\lambda_m)(t^2 + \alpha_1 t + \beta_1) \dots (t^2 + \alpha_M t + \beta_M)$$

enligt Sats 3.8, där $c \neq 0$, $c \in \mathbb{R}$, och varje $\lambda_j, \alpha_j, \beta_j \in \mathbb{R}$ och $m + M \geq 1$, samt att ekvationen gäller för alla t.

Nu gäller

$$0 = (c_0 I + c_1 T + \dots + c_n T^n) v =$$
$$= c(T - \lambda_1 I) \dots (T - \lambda_m I)(T^2 + \alpha_1 T + \beta_1 I) \dots (T^2 + \alpha_M T + \beta_M I) v,$$

vilket betyder att

$T - \lambda_j I$ inte är injektiv för åtminstone ett j

<u>eller</u> att

$T^2 + \alpha_j T + \beta_j I$ inte är injektiv för åtminstone ett j.

Om $T - \lambda_j I$ inte är injektiv för åtminstone ett j så har T ett egenvärde och således ett 1-dimensionellt invariant underrum.

Låt oss betrakta den andra möjligheten. Med andra ord, antag att $T^2 + \alpha_j T + \beta_j I$ inte är injektiv för åtminstone ett j. Då finns ett $0 \neq v \in V$ med

$$T^2 v + \alpha_j T v + \beta_j v = 0.$$

Vi avslutar beviset genom att visa att $[v, Tv]$, som har dimensionen 1 eller 2, är invariant under T.

Betrakta ett tyguckt element i $[v, Tv]$, d.v.s.

$$av + bTv, \text{ där } a, b \in \mathbb{R}.$$

Nu gäller

$$T(av + bTv) = aTv + bT^2v$$

$$= aTv + b(-\alpha_j Tv - \beta_j v)$$

$$= aTv - b\alpha_j Tv - b\beta_j v$$

$$= (a - b\alpha_j)Tv - b\beta_j v,$$

vilket visar att $T(av + bTv) \in [v, Tv]$. Sålедes är $[v, Tv]$ invariant under T.

Vi har sett ett exempel på en operator från $\mathbb{R}^2$ till $\mathbb{R}^2$ som saknar egenvärden. Följande resultat visar att sådana exempel inte existerar för $\mathbb{R}^3$.

<u>Sats 4.10</u> Varje operator från ett uddadimensionellt reellt vektorrum till sig självt har ett egenvärde.

<u>**Bevis:**</u> Antag att V är ett reellt vektorrum med udda dimension. Vi visar att varje operator från V till V har ett egenvärde genom induktion över $\dim V$. Klart att det önskade resultatet gäller om $\dim V = 1$.

Antag att dim V är ett udda tal och större än 1. 180

Genom att använda induktion, kan vi anta att det önskade resultatet gäller för alla operatorer på alla reella vektorrum med dimensionen lika med dim $V-2$.

Antag $T \in \mathcal{L}(V)$. Vi bör visa att T har ett egenvärde. Om T inte har något egenvärde, så ger Sats 4.9 att det finns ett 2-dimensionellt underrum U av V som är <u>invariant</u> under T. Enligt Sats 1.14 existerar ett underrum W av V med

$$V = U \oplus W.$$

Nu gäller $\dim V = \underbrace{\dim U}_{=2} + \dim W$, varför $\dim W = \dim V - 2$.

Nu vill vi använda induktionsantagandet för $T|_W$. Men W behöver inte vara invariant under T, d.v.s. $T|_W$ behöver inte vara en operator från W till W.

Därför definierar vi $S \in \mathcal{L}(W)$ genom

$$Sw = P_{W,U}(Tw) \text{ för } w \in W.$$

Enligt induktionsantagandet har S ett egenvärde λ. Vi skall visa att λ också är ett egenvärde för T.

Låt $0 \neq w \in W$ med $Sw = \lambda w$, d.v.s. $(S-\lambda I)w = 0$. Skulle w vara en egenvektor för T med egenvärdet λ, så skulle vi vara klara.

Typiskt behöver detta inte vara sant.
Därför skall vi söka en egenvektor till T i underrummet $U + [w]$ i V. Tag en typisk vektor $u + aw$ i $U + [w]$, där $u \in U$ och $a \in \mathbb{R}$.
Nu gäller

$$(T - \lambda I)(u + aw) = Tu - \lambda u + a(Tw - \lambda w)$$

$$= Tu - \lambda u + a\big(p_{U,W}(Tw) + p_{W,U}(Tw) - \lambda w\big)$$

$$= Tu - \lambda u + a\big(p_{U,W}(Tw) + S(w) - \lambda w\big)$$

$$= Tu - \lambda u + a\,p_{U,W}(Tw).$$

Emedan $T(U) \subseteq U$, följer att $Tu \in U$. Vidare gäller att $\lambda u \in U$ och $a\,p_{U,W}(Tw) \in U$.
Således $(T - \lambda I)(U + [w]) \subseteq U$.

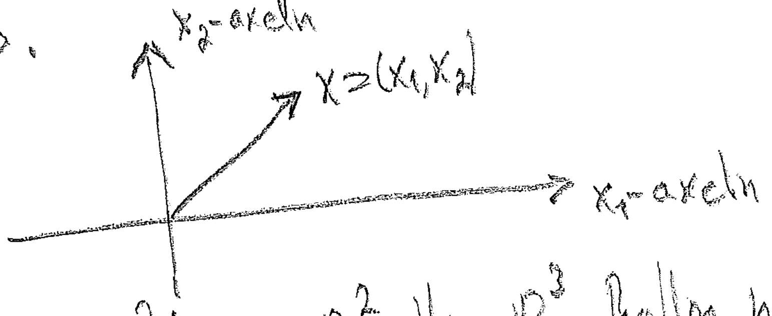
Emedan $\dim(U + [w]) > \dim U$ ger Kor. 2.4 att $(T - \lambda I)|_{U+[w]}$ inte är injektiv, d.v.s.
Det existerar en vektor $x \in U + [w] \subseteq V$, $x \neq 0$, med $(T - \lambda I)x = 0$, d.v.s. $Tx = \lambda x$. Alltså T har ett egenvärde, och satsen är bevisad.

5. Inre produktrum

Då vi definierade ett vektorrum, generaliserade vi den linjära strukturen (addition och skalärmultiplikation) på $\mathbb{R}^2$ och $\mathbb{R}^3$. Vi struntade i vissa viktiga egenskaper såsom längd och vinkel. Dessa ideer finns implanterade i det begrepp som vi nu skall undersöka, dvs. begreppet inre produkt. Vi antar genomgående i fortsättningen att $V \neq \{0\}$ är ett ändligtdimensionellt vektorrum över talkroppen $\mathbb{K}$.

Inre produkt

För att motivera begreppet inre produkt, låt oss tänka på vektorerna i $\mathbb{R}^2$ och $\mathbb{R}^3$ som pilar utgående från origo.



Längden av en vektor i $\mathbb{R}^2$ eller $\mathbb{R}^3$ kallas normen av x och betecknas $\|x\|$. Alltså för $x=(x_1,x_2)\in\mathbb{R}^2$ gäller att $\|x\| = \sqrt{x_1^2+x_2^2}$. På samma sätt gäller för $x=(x_1,x_2,x_3)$ i $\mathbb{R}^3$ att $\|x\| = \sqrt{x_1^2+x_2^2+x_3^2}$. Det är uppenbart hur vi skall generalisera till $\mathbb{R}^n$: Vi definierar normen av $x=(x_1,\dots,x_n)\in\mathbb{R}^n$ genom

$$\|x\| = \sqrt{x_1^2+\dots+x_n^2}.$$

Normen är inte linjär på $\mathbb{R}^n$. För att införa (83)
linearitet introducerar vi skalärprodukten (eller inre produkten) av x och y, som vi betecknar $x \cdot y$, genom

$$x \cdot y = x_1 y_1 + \dots + x_n y_n,$$

där $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ och $y = (y_1, \dots, y_n) \in \mathbb{R}^n$. Notera att skalärprodukten i $\mathbb{R}^n$ är ett reellt tal, inte en vektor. Det är klart att $x \cdot x = \|x\|^2$ för alla $x \in \mathbb{R}^n$. Speciellt gäller att $x \cdot x \geq 0$ för alla $x \in \mathbb{R}^n$ med likhet om och endast om $x = 0$.
Vidare om $y \in \mathbb{R}^n$ är fixerad, är det klart att avbildningen

$$x \in \mathbb{R}^n \longmapsto x \cdot y \in \mathbb{R}$$

är linjär. Dessutom är $x \cdot y = y \cdot x$ för alla $x, y \in \mathbb{R}^n$.

En inre produkt är en generalisering av skalärprodukten. Förrän vi definierar inre produkten på ett vektorrum skall vi titta på det komplexa vektorrummet $\mathbb{C}^n$.

Om $\lambda = a + ib$, där $a, b \in \mathbb{R}$, så definieras absolutbeloppet av λ genom

$$|\lambda| = \sqrt{a^2 + b^2}.$$

och komplexa konjugatet av λ definieras genom $\bar{\lambda} = a - ib$ samt ekvationen

$$|\lambda|^2 = \lambda \bar{\lambda}$$

som sammanbinder dessa två begrepp. För $z = (z_1, \dots, z_n)$ i $\mathbb{C}^n$ definierar vi normen av z genom

$$\|z\| = \sqrt{|z_1|^2 + \dots + |z_n|^2}.$$

Absolutbeloppet behövs eftersom vi vill att $\|z\|$ skall vara ett icke-negativt tal notera att

$$\|z\|^2 = z_1\bar{z}_1 + \ldots\ldots + z_n\bar{z}_n.$$

Vi önskar tänka på $\|z\|^2$ som en inre produkt av z med sig själv såsom vi gjorde i $\mathbb{R}^n$. Således är det naturligt att inre produkten av $w=(w_1, \ldots, w_n)\in\mathbb{C}^n$ med z skall vara lika med

$$w_1\bar{z}_1 + \ldots\ldots + w_n\bar{z}_n.$$

Om vi byter om w och z roll på skulle ovanstående uttryck bytas ut mot dess komplexa konjugat. Med andra ord, det vore naturligt att inre produkten av w och z är lika med komplexa konjugatet av inre produkten av z och w. Efter denna motivering är vi klara att definiera en inre produkt på V som kan vara ett reellt eller komplext vektorrum.

En <u>inre produkt</u> på V är en funktion som avbildar varje ordnat par $(u,v)\in V\times V$ på ett tal $\langle u,v\rangle\in\mathbb{K}$ och har följande egenskaper:

(i) $\langle u,u\rangle \geq 0 \quad \forall u\in V$;

(ii) $\langle u,u\rangle = 0$ om och endast om $u=0$;

(iii) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle \quad \forall u,v,w\in V$;

(iv) $\langle \alpha u,v\rangle = \alpha\langle u,v\rangle \quad \forall \alpha\in\mathbb{K}, \forall u,v\in V$;

(v) $\langle u,v\rangle = \overline{\langle v,u\rangle} \quad \forall u,v\in V$;

Om vi har att göra med reella vektorrum V, så kan den förstnämnda egenskapen formuleras som $\langle u,v\rangle = \langle v,u\rangle$ för alla $u,v \in V$.

Ett inre produktrum är ett vektorrum V försett med en inre produkt på V.
Det viktigaste exemplet på ett inre produktrum är K^n. Vi definierar inre produkten på K^n genom

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\rangle = w_1\bar{z}_1 + \cdots + w_n\bar{z}_n.$$

Det är lätt att se att detta är en inre produkt. Denna inre produkt, som var vår motivering för definitionen av en inre produkt, kallas den euklidiska inre produkten på K^n.
I fortsättningen är K^n, om inget annat sägs, försett med denna inre produkt. Det existerar andra inre produkter på K^n. Till exempel om $c_1, c_2, \ldots c_n$ är positiva tal, så kan vi definiera en inre produkt på K^n genom

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\rangle = w_1\bar{z}_1 c_1 + \cdots + w_n\bar{z}_n c_n.$$

Om $c_1 = \ldots = c_n = 1$, så får vi den euklidiska inre produkten

Som ett annat exempel på ett inre produktrum, betrakta vektorrummet $P_n(K)$ av alla polynom med koefficienter i K med gradtal högst n.
Vi definierar en inre produkt på $P_n(K)$ genom

$$\langle p, q\rangle = \int_0^1 p(t)\overline{q(t)}\,dt.$$

Igen, om $\mathbb{K}=\mathbb{R}$, behövs inte det komplexa konjugatet.

Låt oss nu titta på definitionen av inre produkt.
För varje fixerat $v \in V$, definieras funktionen

$$u \in V \longmapsto \langle u, v\rangle \in \mathbb{K}$$

en linjär avbildningen.
Eftersom varje linjär avbildning avbildar 0 på 0, så gäller

$$\langle 0, v\rangle = 0 \text{ för varje } v \in V.$$

Vidare gäller

$$\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle \quad \forall u, v, w \in V.$$

Dessutom gäller att

$$\langle u, \alpha v\rangle = \bar{\alpha}\langle u, v\rangle \quad \forall \alpha \in \mathbb{K}, \forall\, u, v \in V.$$

Notera att i reella vektorrum gäller att $\langle u, \alpha v\rangle = \alpha\langle u, v\rangle$ där $\alpha \in \mathbb{R}$.

Normer

För $u \in V$, definierar vi normen av u, som vi betecknar $\|u\|$, genom

$$\|u\| = \sqrt{\langle u, u\rangle}.$$

Till exempel, om $(z_1, \ldots, z_n) \in \mathbb{K}^n$, så är

$$\|(z_1, \ldots, z_n)\| = \sqrt{|z_1|^2 + \ldots + |z_n|^2}.$$

Som ett annat exempel, om $p \in \mathcal{P}_n(\mathbb{K})$, försedd med inre produkten på sidan 86, följer att

$$\|p\| = \sqrt{\int_0^1 |p(t)|^2 \, dt}.$$

Notera att $\|u\| = 0$ om och endast om $u = 0$, samt att $\|\alpha u\| = |\alpha| \|u\|$ för alla $\alpha \in \mathbb{K}$ och alla $u \in V$.

Två vektorer $u, v \in V$ sägs vara <u>ortogonala</u> om $\langle u, v \rangle = 0$. Notera att $\langle u, v \rangle = 0 \Leftrightarrow \langle v, u \rangle = 0$. Det är klart att $\langle 0, u \rangle = 0 \; \forall u \in V$ och $\langle 0, 0 \rangle = 0$.

För det speciella fall att $V = \mathbb{R}^2$ är följande sats över 2500 år gammal.

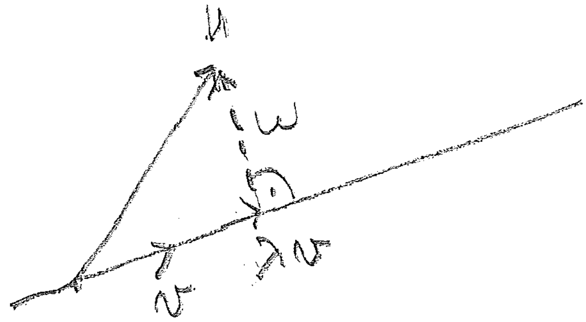
<u>Sats 5.1</u> (Pythagoras sats) Om u och v är ortogonala vektorer i V, så är

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

<u>Bevis:</u> Antag att u och v är ortogonala i V. Då gäller

$$\|u + v\|^2 = \langle u + v, u + v \rangle = \langle u, u + v \rangle + \langle v, u + v \rangle = \langle u, u \rangle + \underbrace{\langle u, v \rangle}_{=0} + \underbrace{\langle v, u \rangle}_{=0} + \langle v, v \rangle = \|u\|^2 + \|v\|^2.$$

Antag att $u, v \in V$. Vi önskar skriva u som en skalär gånger v plus en vektor w som är ortogonal mot v, såsom föreslås av följande figur:



$$u = \lambda v + w$$
och
$$\langle v, w \rangle = 0.$$

Låt $\lambda \in K$. Då gäller

$$u = \lambda v + (u - \lambda v).$$

Vi bör välja λ så att v är ortogonal mot $u - \lambda v$, dvs.

$$0 = \langle u - \lambda v, v \rangle = \langle u, v \rangle - \lambda \|v\|^2.$$

Alltså bör vi välja $\lambda = \langle u, v \rangle / \|v\|^2$, och vi har antagit att $v \neq 0$. Då kan vi skriva

$$u = \frac{\langle u, v \rangle}{\|v\|^2} v + \left(u - \frac{\langle u, v \rangle}{\|v\|^2} v\right).$$

Denna ekvation skall vi använda i beviset till följande sats som ger en av de viktigaste olikheterna i matematiken.

<u>Sats 5.2</u> (Cauchy-Schwarz olikhet)

Om $u, v \in V$ så gäller

$$|\langle u, v \rangle| \leq \|u\| \, \|v\|.$$

Denna olikhet är en likhet om och endast om $u = \alpha v$ för något $\alpha \in K$.

Bevis: Låt $u, v \in V$. Om $v = 0$, så är $|\langle u,v\rangle| = 0 = \|u\|\,\|v\|$. (29
Antag att $v \neq 0$. Betrakta den ortogonala uppdelningen

$$u = \frac{\langle u,v\rangle}{\|v\|^2} v + w$$

där $w = u - \frac{\langle u,v\rangle}{\|v\|^2} v$, och således $\langle w, v\rangle = 0$. Därför fås med Sats 5.1 att

$$\|u\|^2 = \left\| \frac{\langle u,v\rangle}{\|v\|^2} v \right\|^2 + \|w\|^2$$

$$= \frac{|\langle u,v\rangle|^2}{\|v\|^2} + \|w\|^2 \geq \frac{|\langle u,v\rangle|^2}{\|v\|^2},$$

varför $\|u\|\,\|v\| \geq |\langle u,v\rangle|$. Vi ser att likhet gäller om och endast om $\|w\|^2$, dvs. $w = 0$. Men $w = 0$ om och endast om $u = \frac{\langle u,v\rangle}{\|v\|^2} v$, och satsen är bevisad.

Följande resultat kallas triangelolikheten på grund av dess geometriska tolkning, längden av en sida i en triangel är mindre än summan av längden av de två andra sidorna.

Sats 5.3 (Triangelolikheten) Om $u, v \in V$, så är

$$\|u + v\| \leq \|u\| + \|v\|,$$

likhet gäller om och endast om $u = \alpha v$ för något $\alpha \geq 0$.

Bevis: Låt $u, v \in V$. Då gäller

$$\|u + v\|^2 = \langle u + v, u + v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle$$

$= \|u\|^2 + \|v\|^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} = \|u\|^2 + \|v\|^2 + 2\,\mathrm{Re}\langle u,v\rangle$

$\leq \|u\|^2 + \|v\|^2 + 2\,|\langle u,v\rangle| \leq \|u\|^2 + \|v\|^2 + 2\|u\|\,\|v\|$

$= (\|u\| + \|v\|)^2.$

Alltså $\|u+v\| \leq \|u\| + \|v\|$. Vi ser också att likhet gäller om och endast om $\langle u,v\rangle = \|u\|\cdot\|v\| \in \mathbb{R}$.

Om $u = \alpha v$, där $\alpha \geq 0$, följer att

$$\langle u,v\rangle = \alpha\|v\|^2 = \alpha\|v\|\cdot\|v\| = \|u\|\,\|v\|.$$

Omvänt, antag att $\langle u,v\rangle = \|u\|\cdot\|v\|$. Då ger Sats 5.2 att $u = \beta v$, där $\beta \in K$. Nu fås $\beta\|v\|^2 = |\beta|\,\|v\|^2$, dvs. $\beta = |\beta| \geq 0$.

<u>Sats 5.4</u> (Parallellogramidentiteten) Om $u, v \in V$, så är

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

<u>Bevis</u>: Låt $u, v \in V$. Då gäller

$\|u+v\|^2 + \|u-v\|^2 = \langle u+v, u+v\rangle + \langle u-v, u-v\rangle$

$= \|u\|^2 + \|v\|^2 + \langle u,v\rangle + \langle v,u\rangle + \|u\|^2 + \|v\|^2 - \langle u,v\rangle - \langle v,u\rangle$

$= 2(\|u\|^2 + \|v\|^2)$, som vi önskade.

Ortonormerade baser

Vektorerna $e_1, \dots, e_n$ i V kallas <u>ortonormerade</u> om

$$\langle e_j, e_k \rangle = \begin{cases} 0, & \text{om } j \neq k \\ 1, & \text{om } j = k. \end{cases}$$

Till exempel standardbasen i $\mathbb{K}^n$ är ortonormerad.

<u>Sats 5,5</u> Om $e_1, \dots, e_n$ är ortonormerade i V, så gäller

$$\|c_1 e_1 + \dots + c_n e_n\|^2 = |c_1|^2 + \dots + |c_n|^2$$

för alla skalärer $c_1, \dots, c_n$ i $\mathbb{K}$.

<u>Bevis:</u> Eftersom varje e_j har normen 1 följer detta lätt genom upprepad användning av Pythagoras Sats 5.1.

Vi har följande viktiga korollarium.

<u>Korollarium 5.6</u> Om $e_1, \dots, e_n$ är ortonormerade vektorer i V, så är de linjärt oberoende.

<u>Bevis:</u> Låt $c_1, \dots, c_n \in \mathbb{K}$ vara sådana att

$$c_1 e_1 + \dots + c_n e_n = 0.$$

Då ger Sats 5.5 att $|c_1|^2 + \dots + |c_n|^2 = 0$, d.v.s. $c_1 = c_2 = \dots = c_n = 0$, som vi önskade.

En <u>ortonormerad bas</u> i V är ortonormerade vektorer i V som också är en bas i V.

Varje ortonormerad mgd av vektorer i V vars antal är lika med $\dim V$ är automatiskt en ortonormerad bas i V.

<u>Bevis:</u> Enligt Korollarium 5.6 är vektorerna i en sådan mgd linjärt oberoende. Emedan antalet är lika med $\dim V$ måste denna mgd vara en bas för V.

Given en bas $v_1, \dots, v_n$ i V och en vektor $v \in V$, så vet vi att det existerar skalärer $c_1, \dots, c_n \in \mathbb{K}$ med

$$v = c_1 v_1 + \dots + c_n v_n.$$

Men att hitta dessa c_j kan vara svårt. Följande sats visar emellertid att detta är lätt för en ortonormerad bas.

<u>Sats 5.7</u> Antag att $e_1, \dots, e_n$ är en ortonormerad bas i V. Då gäller

$$v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n$$

och

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2$$

för alla $v \in V$.

Bevis: Låt $v \in V$. Då existerar skalärer $c_1, \ldots, c_n \in \mathbb{K}$ så att

$$v = c_1 e_1 + \ldots + c_n e_n.$$

Nu följer $\langle v, e_j \rangle = c_j$, och det första påståendet gäller. Sats 5.5 ger det andra påståendet.

Nu när vi förstår att ortonormerade baser är nyttiga frågar vi oss hur man kan hitta dem? Till exempel, har $P_n(\mathbb{K})$ med inre produkten definierad på sid. 86 en ortonormerad bas? Följande resultat ger ett svar på dessa frågor. Algoritmen som används i följande bevis kallas Gram-Schmidtproceduren. Den ger en metod hur linjärt oberoende vektorer kan göras till ortonormerade vektorer som genererar samma vektorrum som de ursprungliga vektorerna.

Sats 5.8 (Gram-Schmidt) Om $v_1, \ldots, v_n$ är linjärt oberoende vektorer i V, så existerar det en ortonormerad mgd $e_1, \ldots, e_n$ av vektorer i V så att

$$(++) \qquad [v_1, \ldots, v_j] = [e_1, \ldots, e_j] \text{ för } j = 1, \ldots, n$$

Bevis: Antag att $v_1, \ldots, v_n$ är linjärt oberoende vektorer i V. För att konstruera e_i-na börjar vi med att sätta $e_1 = \frac{v_1}{\|v_1\|}$.

Då gäller (**) för $j=1$. Vi skall nu välja $e_2, \ldots, e_n$ induktivt på följande sätt.
Antag att $j > 1$ och att vi valt en ortonormerad mgd $e_1, \ldots, e_{j-1}$, så att

$$[v_1, \ldots, v_{j-1}] = [e_1, \ldots, e_{j-1}].$$

Sätt

(**) $$e_j = \frac{v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}}{\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|}.$$

Notera att $v_j \notin [v_1, \ldots, v_{j-1}]$, ty $v_1, \ldots, v_n$ är linjärt oberoende och därför $v_j \notin [e_1, \ldots, e_{j-1}]$. Således dividerar vi inte med 0 i (**) och följaktligen är e_j väldefinierad. Klart att $\|e_j\| = 1$.
Låt $1 \leq k < j$. Då gäller

$$\langle e_j, e_k\rangle = \frac{\langle v_j, e_k\rangle - \langle v_j, e_k\rangle}{\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|} = 0.$$

Således är $e_1, \ldots, e_j$ en ortonormerad mgd. Från (**) följer att $v_j \in [e_1, \ldots, e_j]$. Emedan $[v_1, \ldots, v_{j-1}] = [e_1, \ldots, e_{j-1}]$ fås härur att

$$[v_1, \ldots, v_j] \subseteq [e_1, \ldots, e_j].$$

Emedan både $v_1, \ldots, v_j$ och $e_1, \ldots, e_j$ är linjärt oberoende mgder, samt $\dim[v_1, \ldots, v_j] = \dim[e_1, \ldots, e_j] = j$, följer att $[v_1, \ldots, v_j] = [e_1, \ldots, e_j]$, och satsen är bevisad.

Nu kan vi ge ett svar på frågan om det alltid existerar någon ortonormerad bas?

<u>Korollarium 5.9</u> Varje ändligtdimensionellt inre produktrum V har en ortonormerad bas.

<u>Bevis:</u> Välj en bas i V. Tillämpa Gram-Schmidt proceduren (Sats 5.8) på denna bas. Då erhåller vi en ortonormerad mgd som är linjärt oberoende (Kor. 5.6), som spänner upp V.

Som vi snart skall se, så räcker det inte alltid att veta att en ortonormerad bas existerar utan det behövs också att en ortonormerad mgd kan utvidgas till en ortonormerad bas.

<u>Korollarium 5.10</u> Varje ortonormerad mgd av vektorer i V kan utvidgas till en ortonormerad bas i V.

<u>Bevis:</u> Antag att $e_1, \dots, e_m$ är en ortonormerad mgd av vektorer i V. Då är $e_1, \dots, e_m$ linjärt oberoende och kan enligt Lemma 1.12 kompletteras till en bas $e_1, \dots, e_m, v_1, \dots, v_n$ i V. Nu tillämpar vi Gram-Schmidt proceduren på denna bas och erhåller en ortonormerad mgd

(X) $$e_1, \dots, e_m, u_1, \dots, u_n.$$

De första m vektorerna förblir oförändrade emedan de redan är ortonormerade. Klart att (x) utgör en ortonormerad bas i V emedan vektorerna i (x) är linjärt oberoende (Kor. 5.6) samt $[e_1, \ldots, e_m, u_1, \ldots, u_n] = V$. Därmed har vi kompletterat $e_1, \ldots, e_m$ till en bas i V.

<u>Korollarium 5.11</u> Antag att $T \in L(V)$. Om det existerar en bas i V med avseende på vilken T har en uppåt triangulär matrisframställning, så existerar det en ortonormerad bas med denna egenskap.

<u>Bevis:</u> Antag $v_1, \ldots, v_n$ är en bas i V med avseende på vilken T har en uppåt triangulär matrisframställning.

Då är $[v_1, \ldots, v_j]$ invariant under T för varje $j = 1, \ldots, n$, enligt Sats 4.5.

Tillämpa Gram-Schmidt proceduren på $v_1, \ldots, v_n$ och vi erhåller en ortonormerad bas $e_1, \ldots, e_n$ i V och

$$[v_1, \ldots, v_j] = [e_1, \ldots, e_j] \text{ för varje } j = 1, \ldots, n,$$

enligt Sats 5.8. Således är $[e_1, \ldots, e_j]$ invariant under T för varje $j = 1, \ldots, n$.

Sats 4.5 ger nu att T har en uppåt triangulär matrisframställning med avseende på den ortonormerade basen $e_1, \ldots, e_n$.

Följande resultat är en viktig tillämpning av ovanstående korollarium.

Korollarium 5.12 Antag att V är ett komplext inre produktrum och $T \in L(V)$.
Då existerar det en ortonormerad bas i V med avseende på vilken T har en uppåt triangulär matrisframställning.

Bevis: Detta följer omedelbart från Sats 4.8 och Korollarium 5.11.

Ortogonala projektioner

Om U är en delmängd av inre produktrummet V, så definieras ortogonala komplementet till U, som vi betecknar $U^\perp$, att vara mängden av alla vektorer i V som är ortogonala mot varje vektor i U, dvs.

$$U^\perp = \{v \in V : \langle v, u \rangle = 0 \ \forall u \in U\}.$$

Notera att $U^\perp$ alltid är ett underrum av V, att $V^\perp = \{0\}$ och $\{0\}^\perp = V$. Vidare gäller att $U_1 \subseteq U_2$ implicerar att $U_2^\perp \subseteq U_1^\perp$.

Följande sats visar att i ett inre produktrum gäller att varje underrum leder till en naturlig direkt uppdelning av hela rummet.

<u>Sats 5.13</u> Om U är ett underrum av V, så gäller

$$V = U \oplus U^{\perp}.$$

<u>Bevis:</u> Låt $e_1, \ldots, e_k$ vara en ortonormerad bas i U. För $v \in V$ gäller

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_k\rangle e_k}_{:= u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_k\rangle e_k}_{:= w}$$

Klart att $u \in U$. För varje j gäller

$$\langle w, e_j\rangle = \langle v, e_j\rangle - \langle v, e_j\rangle = 0.$$

Alltså $w \in [e_1, \ldots, e_k]^{\perp} = U^{\perp}$. Således kan vi skriva

$$v = u + w, \text{ där } u \in U \text{ och } w \in U^{\perp}.$$

Om $v \in U \cap U^{\perp}$, så gäller att $\langle v, v\rangle = 0$, d.v.s. $v = 0$. Följaktligen

$$V = U \oplus U^{\perp}.$$

Vi har följande följd av Sats 5.13.

<u>Korollarium 5.14</u> Om U är ett underrum av V, så är $U = (U^{\perp})^{\perp}$.

Bevis: Tag $u \in U$. Då är $\langle u, v \rangle = 0$ $\forall v \in U^\perp$, dvs. $u \in (U^\perp)^\perp$.

Låt nu $v \in (U^\perp)^\perp$. Enligt Sats 5.13 gäller

$$v = u + w, \text{ där } u \in U \text{ och } w \in U^\perp.$$

Vi får $\langle v, w \rangle = \langle u, w \rangle + \|w\|^2 = \|w\|^2$. Men då $\langle v, w \rangle = 0$, följer att $\|w\| = 0$, dvs. $w = 0$. Alltså $v = u \in U$, varför $(U^\perp)^\perp \subseteq U$, och vi är klara. □

Antag att U är ett underrum av V. Uppdelningen $V = U \oplus U^\perp$ betyder att varje vektor $v \in V$ kan entydigt skrivas i formen

$$v = u + w, \text{ där } u \in U, w \in U^\perp.$$

Vi skall använda denna uppdelning att definiera en avbildning på V, som vi betecknar P_U och kallar ortogonala projektionen av V på U. För $v \in V$ definierar vi

$$P_U(v) = u$$

i uppdelningen ovan. Med våra tidigare beteckningar har vi att $P_U = P_{U,U^\perp}$. Notera att

$$v - P_U v = w$$

i uppdelningen ovan. Således är $P_U v$ en entydig vektor i U sådan att $v - P_U v \in U^\perp$.

Vi skall nu visa att $P_U : V \to V$ är linjär och har följande egenskaper:

(ii) $p_U(V) = U$

(iii) $N(p_U) = U^{\perp}$

(iiii) $p_U^2 = p_U p_U = p_U$

(iv) $\|p_U v\| \leq \|v\| \quad \forall v \in V$

(v) Om $e_1, \dots, e_n$ är en ortonormerad bas i U, då är $\quad p_U v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n \quad \forall v \in V$.

p_U är linjär, ty antag att $v_1 = u_1 + w_1$ och $v_2 = u_2 + w_2$ där $u_1, u_2 \in U$ och $w_1, w_2 \in U^{\perp}$. Då gäller för varje $\alpha \in K$,

$$\alpha v_1 + v_2 = (\alpha u_1 + u_2) + (\alpha w_1 + w_2)$$

Då följer att $\alpha u_1 + u_2 \in U$ och $\alpha w_1 + w_2 \in U^{\perp}$, emedan U och $U^{\perp}$ är underrum av V. Alltså

$$p_U(\alpha v_1 + v_2) = \alpha u_1 + u_2 = \alpha p_U(v_1) + p_U(v_2),$$

dvs. p_U är linjär.

Klart att $p_U(V) = U$. Vidare gäller att $v \in N(p_U)$, dvs. $p_U(v) = 0$ om och endast om $v = w \in U^{\perp}$.

Alltså $N(p_U) = U^{\perp}$.

För godtyckligt $v \in V$ gäller att $p_U(v) \in U$, varav följer att $p_U(p_U(v)) = p_U(v)$. Således är $p_U^2 = p_U$.

Ur uppdelningen $v = p_U v + w$, då $\langle p_U v, w \rangle = 0$, följer att

$$\|v\|^2 = \|p_U v\|^2 + \|w\|^2 \geq \|p_U v\|^2 \quad \forall v \in V.$$

Därför är
$$\|p_U(v)\| \leq \|v\| \quad \forall v \in V.$$

Låt $v \in V$. Enligt Kor 5.10 kan $e_1, \dots, e_n$ i U utvidgas till en ortonormerad bas $e_1, \dots, e_m$ i V. Enligt Sats 5.7 gäller

$$v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_m \rangle e_m.$$

Nu följer

$$p_U v = \langle v, e_1 \rangle p_U e_1 + \dots + \langle v, e_m \rangle p_U e_m.$$

Eftersom $e_{n+1}, \dots, e_m \in U^\perp$ följer att $p_U e_j = 0$ för $j = n+1, \dots, m$ och $p_U e_i = e_i$ för $i = 1, \dots, n$. Därmed följer att

$$p_U v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n.$$

Följande problem uppstår ofta: Givet ett underrum U av V och en vektor $v \in V$. Sök den vektor $u \in U$ så att $\|v - u\|$ är så liten som möjligt.

För att lösa detta problem, betraktar vi en typisk vektor $u \in U$. Då gäller

$$\|v - u\|^2 = \|v - p_U v + p_U v - u\|^2$$

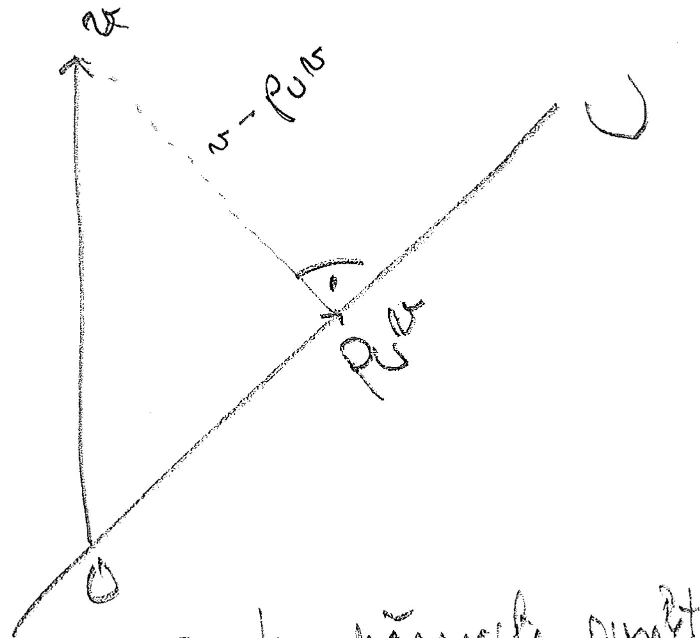
Eftersom $p_U v \in U$ och $p_U v - u \in U$ och uppdelningen $v - p_U v = w \in U^\perp$, ger Pythagoras sats att

$$\|v - u\|^2 = \|v - p_U v\|^2 + \|p_U v - u\|^2.$$

I denna relation är $v \in V$ och U fixerade, så vi har ingen kontroll över termen $\|v - P_U v\|^2$. Men vi kan välja $u \in U$ att vara vilken punkt som helst i U.

För att göra $\|v - u\|$ så liten som möjligt bör vi välja $u \in U$, så att $\|P_U v - u\| = 0$.

Således är $u = P_U v$ det entydiga valet av $u \in U$ för att göra $\|v - u\|$ så liten som möjligt.



$P_U v$ är den närmaste punkten i U till v.

Vidare gör vi att avståndet från v till U, som är definierat som det minsta värde av $\|v - u\|$ då u genomlöper U, är $\|v - P_U v\|$

Notera att om vi har en ortonormerad bas $e_1, \dots, e_n$ i U så är den närmaste punkten i U till v lika med

$$\langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n$$

enligt (V) på sidan 100.

Exempel Bestäm $f \in P_3(\mathbb{R})$ sådan att $f(0)=0$, $f'(0)=0$ och

$$\int_0^1 |2-3t-f(t)|^2\,dt$$

är så liten som möjligt.

Sätt $U=\{f\in P_3(\mathbb{R}) : f(0)=0,\ f'(0)=0\}$. Då är U ett underrum av $P_3(\mathbb{R})$.
Låt $f\in P_3(\mathbb{R})$. Då är $f(t)=c_0+c_1t+c_2t^2+c_3t^3$ och $f'(t)=c_1+2c_2t+3c_3t^2$. Alltså om $f\in U$, så är $f(t)=c_2t^2+c_3t^3$.
Vidare är $p(t)=2-3t\in P_3(\mathbb{R})$.
Klart att t^2, t^3 är en bas i U.
Inre produkten i $P_3(\mathbb{R})$ är definierad genom

$$\langle p,q\rangle=\int_0^1 p(t)q(t)\,dt.$$

Eftersom $\|t^2\|^2=\int_0^1 t^4\,dt=1/5$, sätter vi $e_1(t)=\sqrt{5}\,t^2$. I nästa steg sätter vi

$$f_2(t)=t^3-\langle e_1,t^3\rangle\sqrt{5}\,t^2 = t^3-5t^2\int_0^1 t^5\,dt = t^3-\frac{5}{6}t^2.$$

Emedan $\|t^3-\frac{5}{6}t^2\|^2 = 1/252$, fås att $e_2(t)=6\sqrt{7}\,(t^3-\frac{5}{6}t^2)$.
Vi söker $g(t)=P_U(2-3t)$. Nu gäller

$$\langle 2-3t, e_1(t)\rangle = -\frac{\sqrt{5}}{12}$$

och

$$\langle 2-3t, e_2(t)\rangle=\frac{11\sqrt{7}}{60}.$$

Alltså
$$g(t)= -\frac{5}{12}t^2+\frac{11\sqrt{7}}{60}\,6\sqrt{7}\,\left(t^3-\frac{5}{6}t^2\right)$$
$$=\frac{1}{2}\left(12t^2+\frac{77}{5}t^3\right).$$