4. Jordan-féle normálalak

4.1 λ-mátrixok; kanonikus diagonális alak, invariáns osztók, elemi osztók

4.1.1 Definíció: $M_n(K[\lambda])$ elemeit (K tetszőleges test) λ-mátrixoknak hívjuk.

Legyenek $M_n(K[\lambda])$ elemei felett az „elemi transzformációk" a következők:
- a mátrix egy sorát vagy oszlopát megszorozzuk egy K-beli (nem nulla) elemmel,
- a mátrix egy sorához (oszlopához) hozzáadjuk egy másik sor (oszlop) $c(\lambda)$-szorosát, ahol $c(\lambda)\in K[\lambda]$.

4.1.2 Definíció: az $\mathbf{F},\mathbf{G}$ λ-mátrixok ekvivalensek ($\mathbf{F}\sim\mathbf{G}$), ha véges sok elemi transzformációval egymásba vihetőek. Ez ekvivalencia-reláció lesz. Ugyanis nyilván reflexív és tranzitív, a szimmetria pedig abból következik, hogy elemi transzformációk inverzei is elemi transzformációk.

4.1.3 Megjegyzés: a sor- és oszlopcsere előáll elemi transzformációk kompozíciójaként. Például az u és v sorok cseréje: $(u,v)\to(u,v+u)\to(-v,v+u)\to(-v,u)\to(v,u)$.

4.1.4 Definíció: egy λ-mátrix kanonikus diagonális alakú, ha főátlón kívüli elemei 0-k, az átlón lévő $a_1(\lambda),a_2(\lambda),\dots,a_n(\lambda)$ elemek mindegyike vagy azonosan 0 polinom, vagy 1 a főegyütthatója, továbbá minden átlóelem osztja a következőt (mint polinom).

4.1.5 Tétel: minden ekvivalencia-osztályban legalább egy kanonikus diagonális elem van.

Bizonyítás: $n=1$ esetén az állítás triviális.

Legyen most $n>1$ és legyen $\mathbf{A}$ egy $n\times n$-es λ-mátrix. Ha minden eleme 0, akkor kanonikus diagonális. Ha nem, akkor vegyünk a vele ekvivalens mátrixok közül egy olyat, amiben a bal felső elem a lehető legalacsonyabb fokú, de nem azonosan 0. Legyen ez a mátrix $\mathbf{B}$, bal felső eleme $f(\lambda)$ és legyen az első oszlop egy másik eleme $g(\lambda)$. Osszuk el $g(\lambda)$-t maradékosan $f(\lambda)$-val (megtehetjük, mert K test): $g(\lambda)=q(\lambda)\cdot f(\lambda)+r(\lambda)$, ahol r egy f-nél alacsonyabb fokú vagy azonosan 0 polinom. Vonjuk le g sorából az első sor q-szorosát és cseréljük fel ezt a sort az elsővel. Így a bal felső elem r lesz, a kapott mátrix pedig ekvivalens $\mathbf{B}$-vel, mert lépéseink előállnak elemi transzformációk kompozícióiként. f minimalitása miatt r nem lehet alacsonyabb fokú nem 0 polinom, tehát $r(\lambda)\equiv 0$. Eszerint $\mathbf{B}$-ben az első oszlop - és hasonlóan az első sor - minden eleme többszöröse f-nek. Az első sor ill. az első oszlop megfelelő többszörösét levonva a többi sorból ill. oszlopból az első sor ill. oszlop többi eleme 0 lesz és a kapott $\mathbf{C}$ mátrix is ekvivalens lesz $\mathbf{A}$-val. Legyen az első sor és oszlop elhagyásával $\mathbf{C}$-ben keletkező részmátrixot $\mathbf{D}$, egy tetszőleges eleme $h(\lambda)$. Ennek a sorát hozzáadva az első sorhoz ($\mathbf{C}$-ben) az első sorban megjelenik $h(\lambda)$, tehát a fenti meggondolás alapján h is többszöröse f-nek. Eszerint $\mathbf{D}$ minden eleme többszöröse f-nek.

Ezután vegyük a $\mathbf{D}$ mátrixot. Ennek elemi transzformációit tekinthetjük $\mathbf{C}$ elemi transzformációiként is, mert a $\mathbf{C}$ első oszlopában és első sorában lévő csak 0 elemek vannak (a bal felső elem kivételével), így $\mathbf{D}$ elemi transzformációi nem változtatják meg $\mathbf{C}$ $\mathbf{D}$-n kívüli elemeit. $\mathbf{D}$-re alkalmazva az indukciós feltevést azt kapjuk, hogy ekvivalens egy $\mathbf{E}$ mátrixszal, amely kanonikus diagonális. Beírva ezt $\mathbf{D}$ helyére továbbra is $\mathbf{A}$-val ekvivalens mátrixot kapunk.

A $\mathbf{D}$ mátrix minden eleme f többszöröse volt és elemi transzformációkkal ezt nem tudjuk elrontani, tehát $\mathbf{E}$ elemei is többszörösei. Az első sort leosztva f főegyütthatójával kanonikus diagonális mátrixot kapunk, amely ekvivalens $\mathbf{A}$-val.

Most azt akarjuk belátni, hogy ha két kanonikus diagonális mátrix ekvivalens, akkor azonos.

4.1.6 Definíció: egy λ-mátrix k-adrendű minorja alatt egy $k\times k$-as részmátrixának determinánsát értjük (ez egy λ-polinom). $D_k(\lambda)$-val jelöljük egy λ-mátrix összes k-adrendű minorjának legnagyobb közös osztóját. (Test feletti polinomokra ez létezik.) Az egyértelműség kedvéért $D_k(\lambda)\neq 0$ esetén a főegyütthatót válasszuk 1-nek.

Megjegyzés: $D_k(\lambda)=0 \Leftrightarrow k$ nagyobb a λ-mátrix rangjánál. Ugyanis pontosan ekkor lesz minden $k\times k$-as minor 0.

4.1.7 Lemma: ha két λ-mátrix ekvivalens, akkor mindkettőnél ugyanazokat a $D_k(\lambda)$ polinomokat kapjuk.

Bizonyítás: elég belátni, hogy elemi transzformációk nem változtatják meg $D_k(\lambda)$-t, amihez elég, hogy egy elemi transzformáció utáni új $D_k^*(\lambda)$ többszöröse az eredeti $D_k(\lambda)$-nak. (Az inverz transzformáció is elemi transzformáció, így $D_k(\lambda)$ is többszöröse lesz, $D_k^*(\lambda)$-nak, főegyütthatóik egyenlőek. Ekkor a két polinom azonos.)

Legyen a λ-mátrix **F**. Sor vagy oszlop nem nulla konstanssal való szorzása egy minort legrosszabb esetben konstanssal szoroz, de az osztóin nem változtat. Az i-edik sorhoz a j-edik sor $c(\lambda)$-szorosát adva egy k-adrendű minorral az alábbi három dolog történhet:

– a minorhoz tartozó **A** részmátrixban nem szerepel a j-edik sor; ekkor a minor nem változik, osztható marad $D_k(\lambda)$-val.

– mindkét sor szerepel **A**-ban; ekkor **A**-ban egy sor $c(\lambda)$-szorosát adtuk egy másikhoz, amitől a determinánsa – a minor – nem változik.

– csak a megváltozott sor szerepel **A**-ban. Legyen ekkor $\mathbf{A}_1$ az eredeti minorhoz tartozó mátrix (ennek determinánsa osztható $D_k(\lambda)$-val), $\mathbf{A}_2$ pedig az a mátrix, amit úgy kapunk, hogy ebben az **F** j-edik sorának megfelelő sor helyére az i-edik sornak megfelelő sort írjuk. $\mathbf{A}_2$ a sorok sorrendjétől eltekintve **F** egy minorjának mátrixa lesz, tehát $c(\lambda)\cdot\mathbf{det}(\mathbf{A}_2)$ osztható lesz $D_k(\lambda)$-val. Az $\mathbf{A},\mathbf{A}_1,\mathbf{A}_2$ mátrixok egy soruktól eltekintve azonosak, abban a sorban pedig az **A**-ban lévő sor az $\mathbf{A}_2$-ben lévő sor $c(\lambda)$-szorosának és az $\mathbf{A}_1$-ben lévő sornak az összege, azaz $\mathbf{det}(\mathbf{A})=\mathbf{det}(\mathbf{A}_1)+c(\lambda)\cdot\mathbf{det}(\mathbf{A}_2)$. A jobb oldal többszöröse $D_k(\lambda)$-nak, tehát a bal is. Ezt akartuk belátni.

4.1.8 Következmény: ha két kanonikus diagonális mátrix ekvivalens, akkor azonos, azaz minden ekvivalencia-osztályban legfeljebb egy kanonikus diagonális elem van.

Bizonyítás: azt látjuk be, hogy ha egy **F** kanonikus diagonális mátrix k-adik átlóeleme $d_k(\lambda)$, akkor **(1)** $d_1(\lambda)=D_1(\lambda)$ és $1<k\leq n$-re $D_k(\lambda)=d_k(\lambda)\cdot D_{k-1}(\lambda)$, továbbá **(2)** $D_{k-1}(\lambda)=0$ esetén $d_k(\lambda)=0$.

(1) Mivel **F** diagonális, egy nullától különböző k-adrendű minorja csak olyan lehet, amelyekben a részmátrix sorait és oszlopait szimmetrikusan választottuk. Ezek diagonális mátrixok, melyek átlóelemei az **F** átlóelemei közül valók. Egy ilyen részmátrix determinánsa a kiválasztott elemek szorzata. Ha épp az első k oszlopot és sort választottuk ki, akkor ez a szorzat $\prod_{i=1}^{k} d_i(\lambda)$, különben ennek többszöröse, hiszen minden $d_i(\lambda)$ osztja az utána következőeket. Így $D_k(\lambda)=\prod_{i=1}^{k} d_i(\lambda)$, ebből következik **(1)**.

(2) Ha $D_{k-1}(\lambda)=0$, akkor a $D_k(\lambda)=\prod_{i=1}^{k} d_i(\lambda)$ képletből $K[\lambda]$ nullosztómentessége miatt valamely $i\leq(k-1)$-re $d_i(\lambda)=0$. $d_i(\lambda)$ osztja $d_k(\lambda)$-t, azaz $d_k(\lambda)=0$.

4.1.9 Definíció: az $\mathbf{A}\in M_n(K[\lambda])$ λ-mátrix invariáns osztói – $d_1(\lambda),d_2(\lambda),\ldots,d_n(\lambda)$ – az **A**-val ekvivalens kanonikus diagonális mátrix átlóelemei. Nyilván $\mathbf{det}(\mathbf{A})=c\cdot D_n(\lambda)=c\cdot\prod_{k=1}^{n} d_k(\lambda)$, ahol c a $\mathbf{det}(\mathbf{A})\in K[\lambda]$ polinom főegyütthatója.

4.1.10 Állítás: egy λ-mátrixnak pontosan akkor van inverze, ha determinánsa nem nulla konstans.

Bizonyítás: a determinánsok szorzási szabálya értelmében eleve csak olyan mátrixok lehet inverze, melynek determinánsának van inverze $K[\lambda]$ felett, tehát nem nulla konstans. Ha viszont $\mathbf{det}(\mathbf{U})=U\in K\setminus\{0\}$, akkor vegyük az **U** előjeles aldeterminánsaiból képzett mátrix transzponáltját, $\mathbf{U}^*$-ot. Akár a tisztességes mátrixoknál, itt is $\mathbf{U}\cdot\mathbf{U}^*=\mathbf{I}\cdot U$. Eszerint $\mathbf{U}^*\cdot U^{-1}$ épp a keresett inverz mátrix. Invertálható mátrixok szorzata is invertálható.

4.1.11 Lemma: minden elemi transzformációhoz létezik olyan invertálható λ-mátrix, amellyel egy tetszőleges (persze azonos méretű) λ-mátrixot valamelyik oldalról megszorozva éppen azt teszi, amit a transzformáció.

Bizonyítás: Vegyük azt a diagonális mátrixot, melyben az i-edik átlóelem $\alpha\in K\setminus\{0\}$, a többi 1. Ez egy invertálható mátrix lesz, amellyel balról szorozva egy λ-mátrixot az i-edik sora α-szorosára változik, a többi változatlan; jobbról szorozva az i-edik oszlop lesz az eredeti érték α-szorosa és a többi változatlan. Legyen most $i\neq j$, a **T** mátrixban pedig legyen minden átlóelem 1, $t_{ij}=c(\lambda)$, a többi elem pedig 0. Ez is invertálható mátrix, hiszen determinánsa 1. Könnyen ellenőrizhetően az $\mathbf{A}\mapsto\mathbf{T}\cdot\mathbf{A}$ transzformáció az **A** mátrix j-edik sorához az i-edik sor $c(\lambda)$-szorosát adja, $\mathbf{A}\mapsto\mathbf{A}\cdot\mathbf{T}^T$ pedig a j-edik oszlophoz adja az i-edik oszlop $c(\lambda)$-szorosát. Ezzel minden elemi transzformációhoz találtunk megfelelő invertálható mátrixot. Ha még a másik oldalról szorzunk az egységmátrixszal, akkor minden elemi transzformációt le tudunk írni $\mathbf{A}\mapsto\mathbf{T}_1\cdot\mathbf{A}\cdot\mathbf{T}_2$ alakban, ahol $\mathbf{T}_1,\mathbf{T}_2$ invertálható λ-mátrixok.

4.1.12 Állítás: az $\mathbf{A}$ és $\mathbf{B}$ λ-mátrixok pontosan akkor ekvivalensek, ha léteznek olyan $\mathbf{P},\mathbf{Q}$ invertálható λ-mátrixok, melyekre $\mathbf{B}=\mathbf{PAQ}$.

Bizonyítás: legyenek először $\mathbf{A}$ és $\mathbf{B}$ ekvivalensek, az $\mathbf{A}\mapsto\mathbf{B}$ elemi transzformáció-sorozat legyen $(t_k)_{k=1}^{s}$. Minden t_k-hoz találhatunk **4.1.11** alapján olyan $\mathbf{P}_k,\mathbf{Q}_k$ invertálható mátrixokat, hogy t_k épp az $\mathbf{M}\mapsto\mathbf{P}_k\cdot\mathbf{A}\cdot\mathbf{Q}_k$ transzformáció. Ekkor – hiszen a mátrixszorzás asszociatív – $\mathbf{B}=\mathbf{A}t_1t_2\ldots t_s=\mathbf{P}_s\ldots\mathbf{P}_2\mathbf{P}_1\mathbf{A}\mathbf{Q}_1\ldots\mathbf{Q}_s=(\prod\mathbf{P}_k)\cdot\mathbf{A}\cdot(\prod\mathbf{Q}_k)$. $\mathbf{P}=(\prod\mathbf{P}_k),\mathbf{Q}=(\prod\mathbf{Q}_k)$ éppen a keresett invertálható λ-mátrixok.

Legyen most $\mathbf{B}=\mathbf{PAQ}$. Invertálható mátrixra $D_n(\lambda)=1$, mert a determinánsa nem 0 konstans. Eszerint minden invariáns osztója 1, így ekvivalens az egységmátrixszal. A fenti módon írjuk át $\mathbf{P}$-t $(\prod\mathbf{S}_k)\cdot\mathbf{I}\cdot(\prod\mathbf{T}_k)$ és $\mathbf{Q}$-t $(\prod\mathbf{U}_m)\cdot\mathbf{I}\cdot(\prod\mathbf{V}_m)$ alakra, ahol ezek elemi transzformációk mátrixai. Ekkor $(\prod\mathbf{S}_k)(\prod\mathbf{T}_k)\cdot\mathbf{A}\cdot(\prod\mathbf{U}_m)(\prod\mathbf{V}_m)$ megadja az $\mathbf{A}$-t $\mathbf{B}$-be vivő elemi transzformációkat.

4.1.13 Definíció: vegyünk egy $f(\lambda)\in K[\lambda]$ nem konstans polinomot. Írjuk fel $f(\lambda)=a_0\cdot\prod\varepsilon_k^{n_k}(\lambda)$ alakban, ahol az $\varepsilon_k(\lambda)$ polinomok páronként különböző, K felett irreducibilis, 1 főegyütthatójú polinomok, az n_k kitevők pedig pozitív egész számok. Ekkor $f(\lambda)$ elemi osztói az $\varepsilon_k^{n_k}(\lambda)$ polinomok; $\varepsilon_k^{n_k}(\lambda)$-t-t az $\varepsilon_k(\lambda)$-hoz tartozó elemi osztónak hívjuk. A konstans polinomoknak nincsenek elemi osztói. (Lehetne a 0-nak a konstans 0, a többinek a konstans 1 polinom, ha nagyon akarnánk – nem akarjuk nagyon.) Egy λ-mátrix elemi osztói alatt invariáns osztói elemi osztóinak összességét értjük; ha egy elemi osztó több invariáns osztónál is előfordul, akkor többször is soroljuk fel.

Például ha az invariáns osztók $1,\lambda,\lambda^2(\lambda+1),\lambda^2(\lambda+1)^2,0$ és 0, akkor az elemi osztói $\lambda,\lambda^2,\lambda^2,(\lambda+1),(\lambda+1)^2$.

Az elemi osztók szorzata $D_r(\lambda)$, ahol r a mátrix rangja, hiszen minden nem 0 invariáns osztó egyenlő a belőle kapott elemi osztók szorzatával. Speciálisan, ha a λ-mátrix determinánsa nem 0, akkor a főegyütthatótól eltekintve az elemi osztók szorzata.

4.1.14 Állítás: egy λ-mátrixot az ekvivalencia erejéig egyértelműen meghatároz a mérete (n), a rangja (r) és elemi osztóinak listája.

Bizonyítás: azt kell belátni, hogy az invariáns osztókat egyértelműen meghatározzák. Tudjuk, hogy pontosan az r-nél nagyobb indexű $D_k(\lambda)$-k lesznek 0-k. Pontosan ekkor lesz $d_k(\lambda)=0$. Most már elég $d_1(\lambda),\ldots,d_r(\lambda)$-val foglalkoznunk.

Ha egy $d_k(\lambda)$-nak osztója $\varepsilon^s(\lambda)$, akkor minden továbbinak is osztója a kanonikus alak szerkezete miatt. Vegyük minden (előforduló) $\varepsilon(\lambda)$-ra a listán szereplő hozzá tartozó elemi osztók közül az egyik legmagasabb fokút. Ezek az előző megjegyzés szerint osztják $d_r(\lambda)$-t. Másrészt $d_r(\lambda)$ minden elemi osztója szerepel a listán, azaz az imént kiválasztott elemi osztók épp $d_r(\lambda)$ elemi osztói, $d_r(\lambda)$ pedig ezek szorzata. Ezeket törölve a listáról ugyanezen a módon kapjuk $d_{r-1}(\lambda)$-t az azonos $\varepsilon(\lambda)$-hoz tartozó maximális kitevőjű elemi osztók szorzataként stb. Így egyértelműen megkaptuk a mátrix kanonikus alakját és épp ez volt a célunk. (Persze egy mátrix is egyértelműen meghatározza a méretét, rangját és az elemi osztóit, tehát ez egy bijekció az ekvivalencia-osztályok és a lehetséges méret-rang-lista hármasok között.)

Most már csak arra vagyunk kiváncsiak, hogy miért jobbak az elemi osztók az invariáns osztóknál. Íme:

4.1.15 Lemma: vegyünk egy diagonális mátrixot, legyenek ennek átlós elemei $f_1(\lambda),f_2(\lambda),\ldots,f_n(\lambda)$. Ekkor az elemi osztók listája az $f_i(\lambda)$ polinomok elemi osztói listájának összefűzése.

Bizonyítás: nyilván elég az állítást olyan mátrixokra belátni, ahol az átló egyik eleme sem 0. (A 0 elemek sehol sem adnak elemi osztókat.) Tekintsünk egy $\varepsilon(\lambda)$ irreducibilis polinomot, szerepeljen ez $f_i(\lambda)$ felbontásában az s_i-edik hatványon. Az átlóelemek cserélgetése egyik listát sem befolyásolja, azaz feltehetjük, hogy $s_1\le s_2\le\ldots\le s_n$. A nem nulla k-adrendű minorok azok, amelyek az átlóra szimmetrikus részmátrixokhoz tartoznak, ezek determinánsai a kiválasztott átlóelemek szorzatai. Ebben $\varepsilon(\lambda)$ kitevője a megfelelő s_i-k összege. A legnagyobb közös osztóban, $D_k(\lambda)$-ban a legkisebb ilyen összeg lesz $\varepsilon(\lambda)$ kitevője, azaz $\sum_{i=1}^{k}s_i$. Eszerint $d_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)}$-ban a kitevő $\sum_{i=1}^{k}s_i-\sum_{i=1}^{k-1}s_i=s_k$, a kanonikus diagonális alakban tehát rendre ugyanazok az $\varepsilon(\lambda)$-hez tartozó elemi osztók kitevői. Ezt minden előforduló $\varepsilon(\lambda)$-ra kimondva a bizonyítandó állítást kapjuk.

4.1.16 Állítás: blokkdiagonális mátrix (amely néhány kisebb mátrix direkt összege) elemi osztóinak listája azonos a blokkok elemi osztói listájának összefűzésével.

Bizonyítás: ha minden blokkot elemi transzformációkkal kanonikus diagonális alakra hozunk, akkor a blokkok elemi osztói változatlanok. Ezek egyben az egész mátrixnak is elemi transzformációi, így az egész mátrix elemi osztói is megmaradnak. Kapunk egy diagonális mátrixot, így **4.1.15** alapján az egész mátrix elemi osztóinak listája az átlóelemek elemi osztóinak listájával azonos, ami megfelelően csoportosítva megfelel az egyes blokkok elemi osztói listájának.

4.2 λ-polinomok

Mindeddig a λ-mátrixokra mint polinom elemű mátrixokra, azaz $M_n(K[\lambda])$ elemeire gondoltunk. Megtehetjük azonban, hogy $M_n(K)[\lambda]$ elemeinek, mátrixegyütthatós polinomoknak tekintjük őket. (Ilyenkor néha λ-polinomoknak hívjuk őket.)

4.2.1 Definíció: $\mathbf{F}(\lambda)=\sum_{k=0}^{m}\mathbf{A}_k\lambda^{m-k}$ m-edfokú, ha $\mathbf{A}_0\neq 0$. Reguláris m-edfokú, ha $\mathbf{det}(\mathbf{A}_0)\neq 0$. Ekkor a test feletti polinomokhoz hasonlóan:

(1) $deg(\mathbf{F}+\mathbf{G})\leq \mathbf{max}(deg\,\mathbf{F}, deg\,\mathbf{G})$ és $deg\,\mathbf{F}\neq deg\,\mathbf{G}$ esetén egyenlőség áll fenn,

(2) $deg(\mathbf{F}\cdot\mathbf{G})\leq deg\,\mathbf{F}+deg\,\mathbf{G}$ és ha $\mathbf{F},\mathbf{G}$ egyike reguláris, a másik pedig nem 0, akkor egyenlőség áll fenn, ekkor ugyanis nem eshet ki a két főegyüttható szorzata.

A maradékos osztás problémásabb, már csak azért is, mert a szorzás nem kommutatív és eleve csak reguláris polinommal érdemes próbálkozni. Akkor viszont:

4.2.2 Tétel (maradékos osztás): legyenek $\mathbf{A}=\sum_{k=0}^{n}\mathbf{A}_k\lambda^{n-k}, \mathbf{B}=\sum_{k=0}^{m}\mathbf{B}_k\lambda^{m-k}$ λ-polinomok, utóbbi reguláris. Ekkor azon $\mathbf{P},\mathbf{Q},\mathbf{R},\mathbf{S}$ λ-polinomok, melyekre $\mathbf{A}=\mathbf{B}\cdot\mathbf{P}+\mathbf{S}=\mathbf{Q}\cdot\mathbf{B}+\mathbf{R}$ és $\mathbf{R},\mathbf{S}$ $\mathbf{B}$-nél alacsonyabb fokú vagy azonosan 0 λ-polinomok, léteznek és egyértelműek

Bizonyítás: nyilván elég pl. $\mathbf{P}$ és $\mathbf{S}$ létezését és egyértelműségét belátni. Ha $n<k$, akkor $\mathbf{P}=0, \mathbf{S}=\mathbf{A}$ megfelelő. Ha nem, akkor legyen

$$\mathbf{P}_1=\lambda^{n-m}\cdot\mathbf{B}_0^{-1} \text{ és } \mathbf{S}_1=\mathbf{A}-\mathbf{B}\cdot\mathbf{P}_1=\lambda^{n}\cdot(\mathbf{A}_0-\mathbf{B}_0\cdot\mathbf{B}_0^{-1}\cdot\mathbf{A}_0)+\ldots$$

Ez utóbbi alacsonyabb fokú $\mathbf{B}$-nél, vagy azonosan 0. $\mathbf{A}$-t kicserélve $\mathbf{S}_1$-re ugyanígy kaphatjuk $\mathbf{P}_2$-t és $\mathbf{S}_2$-t, stb. A fok mindig csökken, azaz idővel $\mathbf{S}_l=\mathbf{B}\cdot\mathbf{P}_{l+1}+\mathbf{S}_{l+1}$, ahol $\mathbf{S}_{l+1}$ alacsonyabb fokú $\mathbf{A}$-nál, vagy 0. Ekkor $\mathbf{P}=\sum_{k=1}^{l+1}\mathbf{P}_k, \mathbf{S}=\mathbf{S}_{l+1}$ megfelelő lesz. Két különböző azért nem lehet, mert ha lenne, akkor $\mathbf{A}=\mathbf{B}\cdot\mathbf{P}+\mathbf{S}=\mathbf{B}\cdot\mathbf{P}^*+\mathbf{S}^*$ alapján $\mathbf{B}\cdot(\mathbf{P}-\mathbf{P}^*)=\mathbf{S}^*-\mathbf{S}$ lenne. $\mathbf{B}$ reguláris, $\mathbf{P}-\mathbf{P}^*$ foka legalább 0, $\mathbf{S}^*-\mathbf{S}$ foka alacsonyabb $\mathbf{B}$ fokánál, ami ellentmond **4.2.1.2**-nek.

4.2.3 Definíció: az $\mathbf{A},\mathbf{B}$ λ-polinomok (vagy -mátrixok) skalárisan ekvivalensek, ha léteznek olyan $\mathbf{U},\mathbf{V}$ reguláris nulladfokú λ-polinomok (invertálható skalármátrixok), melyekre $\mathbf{A}=\mathbf{U}\cdot\mathbf{B}\cdot\mathbf{V}$.

4.2.4 Állítás: ha az $\mathbf{A},\mathbf{B}$ elsőfokú, reguláris λ-polinomok, melyek (mint λ-mátrixok) ekvivalensek, akkor skalárisan is ekvivalensek.

Bizonyítás: **4.1.12**-ből tudjuk, hogy ⚀ $\mathbf{A}=\mathbf{UBV}$, ahol $\mathbf{U}$ és $\mathbf{V}$ invertálható λ-mátrixok. **4.2.2** alapján vannak olyan $\mathbf{P},\mathbf{Q}$ λ-mátrixok ill. $\mathbf{R},\mathbf{S}$ $\mathbf{A}$-nál alacsonyabb fokú, azaz skalármátrixok, melyekre ⚁ $\mathbf{U}=\mathbf{AP}+\mathbf{S}$, ⚂ $\mathbf{V}=\mathbf{QA}+\mathbf{R}$. Szorozzuk meg ⚀-t $\mathbf{U}^{-1}$-el balról, írjuk $\mathbf{V}$-be ⚂-t és mindezt rendezzük át. Kapjuk, hogy:

$$\mathbf{U}^{-1}\mathbf{A}=\mathbf{BV}=\mathbf{BQA}+\mathbf{BR}$$

⚃ $$(\mathbf{U}^{-1}-\mathbf{BQ})\mathbf{A}=\mathbf{BR}$$

Mivel $\mathbf{A}$ reguláris és $\mathbf{BR}$ legfeljebb elsőfokú, **4.2.1.2** szerint $\mathbf{T}=(\mathbf{U}^{-1}-\mathbf{BQ})$ legfeljebb nulladfokú lehet, azaz skalármátrix. Balról $\mathbf{U}$-val szorozva és átrendezve $\mathbf{I}=\mathbf{UU}^{-1}=\mathbf{UT}+\mathbf{UBQ}$. Jobbról megszorozva ⚀-t $\mathbf{V}^{-1}$-el és a kapott eredményt beírva $\mathbf{UB}$ helyére $\mathbf{I}=\mathbf{UT}+\mathbf{AV}^{-1}\mathbf{Q}$. Beírva $\mathbf{U}$-ba ⚁-t és a kapott kifejezésből $\mathbf{ST}$-t levonva:

$$\mathbf{I}=(\mathbf{AP}+\mathbf{S})\mathbf{T}+\mathbf{AV}^{-1}\mathbf{Q}$$
$$\mathbf{I}-\mathbf{ST}=\mathbf{A}(\mathbf{PT}+\mathbf{V}^{-1}\mathbf{Q})$$

Eme szorzatban $\mathbf{A}$ reguláris elsőfokú, így az eredmény vagy legalább elsőfokú, vagy 0. $(\mathbf{I}-\mathbf{ST})$ viszont skalármátrix, így csak akkor lehetnek egyenlőek, ha mindkét oldal 0. Eszerint $\mathbf{I}=\mathbf{ST}$. Azaz $\mathbf{S}$ nem csupán skalármátrix, invertálható is. Ha a számolásban ⚀$\cdot\mathbf{V}^{-1}$-ből indultunk volna, $\mathbf{U}$ helyére írtuk volna ⚂-t stb., akkor az jött volna ki, hogy $\mathbf{T}^*\mathbf{R}=\mathbf{I}$, így $\mathbf{R}$ is invertálható.

Most már könnyű dolgunk van. ⚃-ben ugyanis kijött $\mathbf{TA}=\mathbf{BR}$, amit balról $\mathbf{S}$-el szorozva $\mathbf{A}=\mathbf{IA}=\mathbf{STA}=\mathbf{SBR}$, így $\mathbf{A}$ és $\mathbf{B}$ skalárisan ekvivalensek ($\mathbf{S},\mathbf{R}$ invertálható skalármátrixok).

4.2.5 Tétel: az $\mathbf{A},\mathbf{B}\in M_n(K)$ mátrixok pontosan akkor hasonlóak, ha a $\lambda\mathbf{I}-\mathbf{A},\lambda\mathbf{I}-\mathbf{B}$ λ-mátrixok ekvivalensek.

Bizonyítás: ha hasonlóak, akkor $\mathbf{A}=\mathbf{P}^{-1}\mathbf{BP}\Rightarrow\lambda\mathbf{I}-\mathbf{A}=\mathbf{P}^{-1}(\lambda\mathbf{I}-\mathbf{B})\mathbf{P}$, a λ-mátrixok skalárisan ekvivalensek. Ha a két λ-mátrix ekvivalens, akkor **4.2.4** szerint skalárisan ekvivalens, azaz $\lambda\mathbf{I}-\mathbf{A}=\mathbf{R}(\lambda\mathbf{I}-\mathbf{B})\mathbf{S}=\lambda\cdot\mathbf{RIS}-\mathbf{RBS}$. Itt minden betű skalármátrix (λ-t kivéve), ezek tehát az együtthatók. Azonos polinomok együtthatói azonosak, tehát egyrészt $\mathbf{RS}=\mathbf{I}\Rightarrow\mathbf{R}=\mathbf{S}^{-1}$, másrészt $\mathbf{A}=\mathbf{RBS}=\mathbf{S}^{-1}\mathbf{BS}$, a két mátrix valóban hasonló.

4.3 A Jordan-féle normálalak

4.3.1 Definíció: egy $\mathbf{N}\in M_n(K)$ mátrix Jordan-féle normálalak, ha az alábbi szerkezetű: k darab blokkból áll, ahol az i-edik blokk $s(i)$ kisebb blokkból áll; az i-edik blokk j-edik részblokkja t_{ij} méretű, főátlójában ϱ_i van, a közvetlenül átló feletti elemek 1-ek, minden más pedig 0 (a ϱ_i-k K páronként különböző elemei).

$$\begin{pmatrix} \varrho & 1 & 0 & \cdots & 0 \\ 0 & \varrho & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \varrho & 1 \\ 0 & \cdots & 0 & 0 & \varrho \end{pmatrix}$$

4.3.2 Állítás: az előbbi jelölésekkel az $\mathbf{N}$ Jordan-féle normálalakra $(\lambda\mathbf{I}-\mathbf{N})$ elemi osztói az $(\lambda-\varrho_i)^{t_{ij}}$ polinomok, ahol $1\le i\le k, 1\le j\le s(i)$. Ebből látszik az is, hogy ha két Jordan-féle normálalak elemi osztóinak listája azonos, ($\Leftrightarrow$ ekvivalensek) akkor csak a blokkok sorrendjében térhetnek el.

Bizonyítás: vegyük az i-edik blokk j-edik részblokkját. Ennek determinánsa az átlóelemek szorzata – az átló alatt minden 0 –, az átlóban t_{ij} darab $(\lambda-\varrho_i)$ van, így épp $(\lambda-\varrho_i)^{t_{ij}}$-t kapunk. Az első oszlop és az utolsó sor elhagyásával egy olyan mátrixot kapunk, amelynek átlójában (-1)-ek vannak, felette csupa 0, determinánsa tehát ±1. Eszerint az utolsóelőtti invariáns osztó 1, hiszen osztja az előbb kapott minort. Minden korábbi invariáns osztó osztja ezt, tehát szintén egy. Az utolsó invariáns osztó $(\lambda-\varrho_i)^{t_{ij}}$, hogy szorzatként kijöjjön a determináns. Ez egy irreducibilis polinom hatványa, tehát személyesen a blokk egyetlen elemi osztója. Az egész mátrix elemi osztóinak listája **4.1.16** alapján ezek összefűzése és épp ezt akartuk bizonyítani.

4.3.3 Tétel: ha K algebrailag zárt test, akkor minden $\mathbf{A}\in M_n(K)$ mátrix hasonló egy $\mathbf{N}$ Jordan-féle normálalakhoz és ez az alak a blokkok sorrendjétől eltekintve egyértelmű.

Bizonyítás: algebrailag zárt test felett pontosan az elsőfokú polinomok irreducibilisek, tehát $(\lambda\mathbf{I}-\mathbf{A})$ elemi osztói elsőfokú polinomokhoz tartoznak. Jelölje az elemi osztókban szereplő irreducibilis polinomokat $(\lambda-\varrho_i)$, ahol $1\le i\le k$; egy ϱ_i-hez tartozó elemi osztók $(\lambda-\varrho_i)^{t_{ij}}$, ahol $1\le j\le s(i)$. Ehhez a listához persze találhatunk egy $\mathbf{N}$ Jordan-féle normálalakot **4.3.2** alapján és ez lényegében egyértelmű. Már csak az kell, hogy ekkor $(\lambda\mathbf{I}-\mathbf{A})$ és $(\lambda\mathbf{I}-\mathbf{N})$ mérete és rangja is azonos, hiszen akkor **4.1.14** szerint ezek ekvivalensek, azaz **4.2.5** szerint $\mathbf{A}$ és $\mathbf{N}$ hasonlóak, ami épp a bizonyítandó állítás. Mindkét λ-mátrix mérete és rangja is azonos determinánsának fokával, a két determináns pedig a megegyezik, mert ugyanazon elemi osztók szorzata és főegyütthatójuk 1.

Megjegyzés: tulajdonképpen azt bizonyítottuk, hogy ha $\chi_{\mathbf{A}}(\lambda)$ lineáris faktorok szorzatára bomlik K felett, akkor van $M_n(K)$-ban $\mathbf{A}$-hoz hasonló $\mathbf{N}$ Jordan-normálalak. Ez a feltétel persze szükséges is, hiszen $\chi_{\mathbf{N}}(\lambda)$ lineáris faktorokra esik és megegyezik $\chi_{\mathbf{A}}(\lambda)$-val.

Megjegyzés: a Jordan-féle normálakra $(\lambda\mathbf{I}-\mathbf{N})$ determinánsa egyrészt $\prod_{i=1}^{k}\prod_{j=1}^{s(i)}(\lambda-\varrho_i)^{t_{ij}}$, másrészt a karakterisztikus polinom. Eszerint a mátrixnak megfelelő lineáris transzformáció sajátértékei a ϱ_i-k, algebrai multiplicitásuk pedig rendre $\sum_{j=1}^{s(i)}t_{ij}$. A geometriai multiplicitások az $s(i)$-k, mert a ϱ_i-hez tartozó sajátaltér éppen az, amit az i-edik blokk részblokkjainak első soraihoz (és oszlopaihoz) tartozó báziselemek feszítenek ki.

Megjegyzés: a minimálpolinom is leolvasható a normálalakról. Ugyanis egy blokkdiagonális mátrixot hatványozva, konstanssal szorozva a blokkok „nem zavarják” egymást, azaz egy polinomba behelyettesítve az egész mátrixot ugyanazt kapjuk, mintha a blokkokat külön-külön behelyettesítenénk és azután raknánk össze. Így az egész mátrix minimálpolinomja a blokkok minimálpolinomjainak legkisebb közös többszöröse. Az i-edik blokk j-edik részblokkjának minimálpolinomja a változatosság kedvéért $(\lambda-\varrho_i)^{t_{ij}}$, a teljes mátrix minimálpolinomja tehát $\prod_{i=1}^{k}(\lambda-\varrho_i)^{\max_{1\le j\le s(i)}t_{ij}}$.

Mivel $\mathbb{C}$ algebrailag zárt, $M_n(\mathbb{C})$-ben minden mátrixnak van Jordan-normálalakja. Sajnálatos módon $\mathbb{R}$ nem zárt algebrailag, így egy valós mátrixnak nem mindig van $M_n(\mathbb{R})$-ben Jordan-normálalakja. Viszont – mivel könnyen le tudjuk írni az $\mathbb{R}$ feletti irreducibilis polinomokat – $M_n(\mathbb{R})$-ben kissé megváltoztatva a normálalak fogalmát már minden mátrixhoz fogunk hasonló normálalakot találni.

4.3.4 Definíció: a $\mathbf{M}\in M_n(\mathbb{R})$ mátrix valós Jordan-normálalak, ha k darab blokkból áll, ahol az m-edik blokk $s(m)$ kisebb blokkból áll; a m-edik blokk j-edik részblokkja az alábbiak valamelyike:

– a $\varrho_m\in\mathbb{R}$ számhoz tartozó t_{mj} méretű „szabványos” blokk, azaz az átlón csupa ϱ_m van, közvetlenül az átló felett csupa 1, másutt csupa 0;

– a $\varrho_m=a_m+b_m\cdot i\in\mathbb{C}\setminus\mathbb{R}$ számhoz tartozó olyan $2t_{mj}$ méretű mátrix, amelynek átlóján néhány $\mathbf{C}_m$ van, közvetlenül fölöttük rendre $\mathbf{I}_2=\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\in M_2(\mathbb{R})$, máshol 0 ($\mathbf{C}_m$ a $\left(\begin{smallmatrix}a_m&b_m\\-b_m&a_m\end{smallmatrix}\right)\in M_2(\mathbb{R})$ mátrixot jelöli).

A ϱ_m-ek páronként különböző, nemnegatív képzetes részű komplex számokat jelölnek.

4.3.5 Állítás: a fenti jelölésekkel az $\mathbf{M}$ valós Jordan-normálalakra $(\lambda\mathbf{I}-\mathbf{M})$ elemi osztói a $(\lambda-\varrho_m)^{t_{mj}}$: $\varrho_m\in\mathbb{R}$ és $\left((\lambda-\varrho_m)(\lambda-\bar{\varrho}_m)\right)^{t_{mj}}$: $\varrho_m\notin\mathbb{R}$ polinomok.

Bizonyítás: vegyük észre, hogy $(\lambda\mathbf{I}-\mathbf{M})$-ben az m-edik blokk j-edik részblokkjának egyetlen elemi osztója $(\lambda-\varrho_m)^{t_{mj}}$ ha $\varrho_m\in\mathbb{R}$ ill. $\left((\lambda-\varrho_m)(\lambda-\bar{\varrho}_m)\right)^{t_{mj}}$ ha $\varrho_m\notin\mathbb{R}$. Ezt a valós esetben tudjuk, hiszen ez ugyanolyan, mint **4.3.2**-ben. A másik eset hasonlóan kiszámolható. $(\lambda\mathbf{I}-\mathbf{M})$ ezen blokkok direkt összege, így elemi osztói épp az általunk megadottak.

4.3.6 Tétel: minden $\mathbf{A}$ valós mátrix hasonló egy és (lényegében) csakis egy valós Jordan-normálalakhoz.

Bizonyítás: tudjuk, hogy $\mathbb{R}[\lambda]$-ban minden irreducibilis polinom lineáris vagy alkalmas pozitív képzetes részű z-re $(\lambda-z)(\lambda-\bar{z})$ alakú. Így $(\lambda\mathbf{I}-\mathbf{A})$ elemi osztóinak listája $(\lambda-\varrho)^t$: $\varrho\in\mathbb{R}$ és $\left((\lambda-(a+bi))(\lambda-(a-bi))\right)^{t^*}$: $a\in\mathbb{R}, b\in R^+$ alakú polinomokból áll. Találhatunk tehát egy $\mathbf{M}$ valós Jordan-normálalakot, amelyre $(\lambda\mathbf{I}-\mathbf{M})$ elemi osztóinak listája ugyanez, ekkor $\mathbf{A}\sim\mathbf{M}$.