Miért működik

A parciális törtekre bontás?

Borbély Gábor

2012. június 7.

Tartalomjegyzék

1. Lineáris algebra formalizmus **2**

2. A feladat kitűzése **3**

3. A LER felépítése **5**

4. A bizonyítás **6**

1. Lineáris algebra formalizmus

Ebben a fejezetben bevezetünk egy vektor-mátrix formalizmust a polinomok kezelésére.

Egy n-ed fokú $p(x)$ polinom értékét felfoghatjuk két, $\mathbb{R}^{n+1}$-beli vektor skalárszorzataként:

$$\begin{aligned} p(x) &= \underline{p} \cdot \underline{x}_n \\ p_0 + p_1 x + p_2 x^2 + \ldots + p_n x^n &= \left\langle \begin{pmatrix} p_0 \\ \vdots \\ p_n \end{pmatrix}, \begin{pmatrix} 1 \\ \vdots \\ x^n \end{pmatrix} \right\rangle . \end{aligned} \tag{1}$$

Ílymódon a polinomot egy $\underline{p} \in \mathbb{R}^{n+1}$ vektor jellemez ($\underline{x}_n$ egy szimbólum, nem adat, az indexe pedig a fokszámot jelöli). Egy vektor egyértelműen megad egy polinomot, de egy polinom nem határoz meg egyértelműen egy vektort, mert ízlés szerint 0 koordinátákat fűzhetünk annak végére. Mosttantól minden $p(x) \in \mathbb{R}_n[x]$ polinomot együtt kezelünk a $\underline{p} \in \mathbb{R}^{n+1}$ vektor megfelelőjével (feltesszük, hogy $\deg(p)$ teljes, vagyis nincs $\underline{p}$ végén 0 elem).

Két polinomot össszeadhatunk az őket reprezentáló vektorok segítségével: $p(x) + q(x) = (\underline{p} + \underline{q}) \cdot \underline{x}$. Ha nem egyenlő fokszámúak, akkor a rövidebbik vektor végére 0-kat fűzünk. Polinomokat szorozni is tudunk. Adott $p(x) \in \mathbb{R}_n[x]$ polinomhoz definiálunk egy szorzó mátrixot, ami a $p(x)$ polinommal való szorzást végzi el.

$$p(x) = p_0 + p_1 x + p_2 x^2 + \ldots + p_n x^n$$

$$\mathrm{D}_l(p) := \underbrace{\begin{pmatrix} p_0 & & & \\ p_1 & p_0 & & \\ \vdots & p_1 & \ddots & \\ p_n & \vdots & \ddots & p_0 \\ & p_n & \ddots & p_1 \\ & & \ddots & \vdots \\ & & & p_n \end{pmatrix}}_{l+1 \text{ darab példánya } \underline{p}\text{-nek}} \in \mathbb{R}^{(n+l+1)\times(l+1)} \tag{2}$$

Így bármilyen $b(x) \in \mathbb{R}_l[x]$ polinomot reprezentáló $\underline{b} \in \mathbb{R}^{l+1}$ vektorra a $\mathrm{D}_l(p) \cdot \underline{b} \in \mathbb{R}^{n+l+1}$ vektor a $p(x) \cdot b(x)$ polinomot reprezentálja. A $\mathrm{D}_l(p)$ mátrix aljára tetszőleges számú 0 sorokat fűzhetünk és annyiszor rakjuk bele a $\underline{p}$ vektort, amekkora polinomot akarunk szorozni vele. Ugyanakkor egy szorzó mátrix első oszlopából kiolvasható, hogy milyen polinommal szoroz, így egy polinom egy mátrixszal is reprezentálható.

Mátrix formalizmussal: legyen N egy nagy természetes szám (aminél nagyobb fokú polinom biztosan nem fordul elő a számolásainkban) és $\mathbb{R}_N[x] \ni p(x) \Leftrightarrow \underline{p} \in \mathbb{R}^{N+1}$ polinom legyen megfeleltetve $\mathrm{D}_N(p)$ mátrixnak.

Lehet négyzetes mátrix-reprezentációt is adni, ezt ${}_N\mathrm{D}(p)$-vel jelöljük.

$$
{}_N\mathrm{D}(p) := \begin{pmatrix} p_0 & & & \\ p_1 & p_0 & & \\ \vdots & & \ddots & \\ p_n & \cdots & & p_0 \end{pmatrix} \in \mathbb{R}^{(N+1)\times(N+1)}
$$

Így a ${}_N\mathrm{D}(p) \in \mathbb{R}^{(N+1)\times(N+1)}$ mátrixok kommutatív gyűrűt alkotnak. Ez nem más, mint az $\mathbb{R}_N[x]$ egy bijektív reprezentációja.

Megjegyezzük azt is, hogy $\mathrm{rank}(D_l(p)) = l+1$ bármilyen nem 0 polinomot is jelöl $\underline{p}$, mivel az $\mathrm{Im}(D_l(p))$ képtér a $\left\langle p(x), x \cdot p(x), \ldots x^l \cdot p(x) \right\rangle$ polinomok lineárisan generált terének felel meg. Valamint ${}_N\mathrm{D}(p)$, ha egy polinomot (mátrixot) balról szoroz, akkor nem csökkenti annak rangját, amennyiben az eredményül kapott polinom foka nem haladja meg az N-et.

2. A feladat kitűzése

Az *algebra alaptételé*ből tudjuk, hogy minden (egy főegyütthatós) $p(x) \in \mathbb{R}_N[x]$ polinom a következő alakban faktorizálható $\mathbb{R}$ felett:

$$
p(x) = \prod_{i=1}^{k'} (x - x_i)^{\nu_i} \cdot \prod_{j=k'+1}^{k} \left((x - x_j)^2 + c_j\right)^{\nu_j} \tag{3}
$$

ahol $\nu_i > 0, \nu_j > 0$, $x_i \in \mathbb{R}, x_j \in \mathbb{R}$, $N = \sum_{i=1}^{k'} \nu_i + \sum_{j=k'+1}^{k} 2\nu_j$, $c_j > 0$ és x_i-k mind különböznek az első csoportban és (x_j, c_j)-k mind különböznek a második csoportban. Hogy értelme legyen parciális törtekre bontani, feltesszük $k \geq 2$-t.

Kicsit egyszerüsítsünk a jelölésen.

$$p(x) = \prod_{i=1}^{k} p_i(x)^{\nu_i} \tag{4}$$

ahol p_i-k $\mathbb{R}$ felett irreducibilis polinomok, mind páronként relatív prímek. Használjuk azt a jelölést is, hogy

$$\mu_i = \begin{cases} \nu_i & \text{ha } p_i \text{ lineáris} \\ 2\nu_i & \text{ha } p_i \text{ másodfokú (nincs valós gyöke)} \end{cases}, \tag{5}$$

így $N = \sum_{i=1}^{k} \mu_i$. Továbbá legyen $P_i(x) := p_i(x)^{\nu_i}$, ezzel

$$p(x) = \prod_{i=1}^{k} P_i(x) \tag{6}$$

ahol P_i-k rendre μ_i-ed fokú polinomok, és semelyik kettő sem tartalmaz közös faktorokat (és feltettük, hogy minden főegyüttható 1).

A parciális törtekre bontást a közetkező alakban fogjuk keresni (azért, hogy egy általános racionális törtfüggvény primitív függvényét megkapjuk):

$$\frac{1}{p(x)} = \sum_{i=1}^{k'} \sum_{\nu=1}^{\nu_i} \frac{a_\nu^i}{p_i(x)^{\nu_i-\nu+1}} + \sum_{j=k'+1}^{k} \sum_{\mu=1}^{\nu_j} \frac{a_{\mu 1}^j + a_{\mu 2}^j \cdot x}{p_j(x)^{\nu_j-\mu+1}} \tag{7}$$

Lényegében egy $\underline{a} \in \mathbb{R}^N$ vektort keresünk, ahol a következő együtthatók olvashatóak ki:

$$\underline{a}^\top = \overbrace{\underbrace{a_1^1 a_2^1 \dots a_{\nu_1}^1}_{\nu_1 \text{ darab}}, \underbrace{a_1^2 a_2^2 \dots a_{\nu_2}^2}_{\nu_2 \text{ darab}}, \dots a_{\nu_k'}^{k'}}^{\text{lineáris tagok}}, \overbrace{\dots, \underbrace{a_{11}^j a_{12}^j \dots a_{\nu_j 1}^j a_{\nu_j 2}^j}_{2\times\nu_j \text{ darab}}, \dots}^{\text{másodfokú tagok}}$$

Szorozzuk be (7)-t $p(x)$-el és használjuk hozzá a következő jelölést:

$$\mathfrak{p}_j(x) := \prod_{\substack{i=1\dots k \\ i\neq j}} P_i(x) \in \mathbb{R}_{N-\mu_j}[x] \tag{8}$$

$$1 = \sum_{i=1}^{k'} \left(\mathfrak{p}_i(x) \sum_{\nu=0}^{\nu_i-1} a_\nu^i p_i(x)^\nu \right) + \sum_{j=k'+1}^{k} \left(\mathfrak{p}_j(x) \sum_{\mu=1}^{\nu_j-1} \left(a_{\mu 1}^j + a_{\mu 2}^j \cdot x \right) p_j(x)^\mu \right) \tag{9}$$

A fenti egyenlőség, már polinomok egyenlősége, ezt fogjuk a fenti formalizmussal egy **lineáris egyenletrendszerré alakítani, majd együtthatómátrixszáról belátjuk, hogy invertálható**.

3. A LER felépítése

Foglalkozzunk először egy darab (multiplicitásokkal rendelkező) faktorral, vagyis egy darab $p_i(x)$ hatványaival.

Tegyük fel, hogy p_i lineáris, ekkor $\sum_{\nu=0}^{\nu_i-1} a_\nu^i p_i(x)^\nu$ nem más, mint az $1, p_i(x), p_i(x)^2, \ldots, p_i(x)^{\nu_i-1}$ polinomok egy lineáris kombinációja. Mátrix formalizmussal:

$$\begin{array}{c}
\begin{array}{ccccc} 1 & p_i & {p_i}^2 & \cdots & {p_i}^{\nu_i-1} \end{array} \\
\left[\begin{array}{ccccc}
1 & * & * & & * \\
0 & 1 & * & & * \\
\vdots & 0 & 1 & & * \\
 & & & \ddots & \vdots \\
0 & 0 & 0 & & 1
\end{array}\right]
\end{array}
\cdot
\left[\begin{array}{c} a_1^i \\ a_2^i \\ \vdots \\ a_{\nu_i}^i \end{array}\right]$$

A főátlóban ν_i darab egyes áll, mert egy $x - x_i$ alakú polinomnak a hatványai rendre $0, 1, \ldots, \nu_i - 1$ fokú, egy főegyütthatójú polinomok. p_i hatványai alatt a p_i polinom hatványait reprezentáló vektorokat értjük, számolhatjuk a fent említett ${}_{\nu_i-1}\mathrm{D}(p_i)$ mátrix hatványozásával is. Látható, hogy a fenti mátrix invertálható $\nu_i \times \nu_i$-es, ami egyben $\mu_i \times \mu_i$-es, mert p_i első fokú volt, lásd (5) jelölést.

Lássuk most a másodfokúakat, a $\sum_{\mu=1}^{\nu_j-1} \left(a_{\mu 1}^j + a_{\mu 2}^j \cdot x\right) p_j(x)^\mu$ polinomot. Ez nem más, mint az $1, x, p_j(x), x \cdot p_j(x), p_j(x)^2, x \cdot p_j(x)^2, \ldots, p_j(x)^{\nu_i-1}, x \cdot p_j(x)^{\nu_i-1}$ polinomok egy lineáris kombinációja.

$$
\begin{array}{ccccccc}
1 & x & p_j & x \cdot p_j & \cdots & {p_j}^{\nu_j-1} & x \cdot {p_j}^{\nu_j-1}
\end{array}
$$

$$
\left[\begin{array}{c|c|c|c|c|c|c}
1 & 0 & * & 0 & & * & 0 \\
0 & 1 & * & * & & * & * \\
\vdots & 0 & 1 & * & & * & * \\
 & \vdots & & 1 & & * & * \\
 & & & & \ddots & & \\
 & & & & & 1 & \\
0 & & & & \cdots & 0 & 1
\end{array}\right]
\cdot
\begin{bmatrix} a^j_{11} \\ a^j_{12} \\ \vdots \\ a^j_{\nu_j 1} \\ a^j_{\nu_j 2} \end{bmatrix}
$$

Itt a blokkok kettesével jönnek, mert a hatványozás során p_j foka kettőt nő, de van közte egy x szorzó beiktatva. Látható hogy a fenti mátrix invertálható és $(2\nu_j) \times (2\nu_j)$-s, ami $\mu_j \times \mu_j$-es, megint használva (5) jelölést.

Nevezzük el a fenti ábrákon látható mátrixokat: $A_i \in \mathrm{SL}_{\mathbb{R}}(\mu_i), i = 1 \dots k$, most már lineáris és másodfokú mind együtt van kezelve. A_i-ket még szorozni kell $\mathfrak{p}_i$-vel, így adják ki a (9) képlet jobb oldalát:

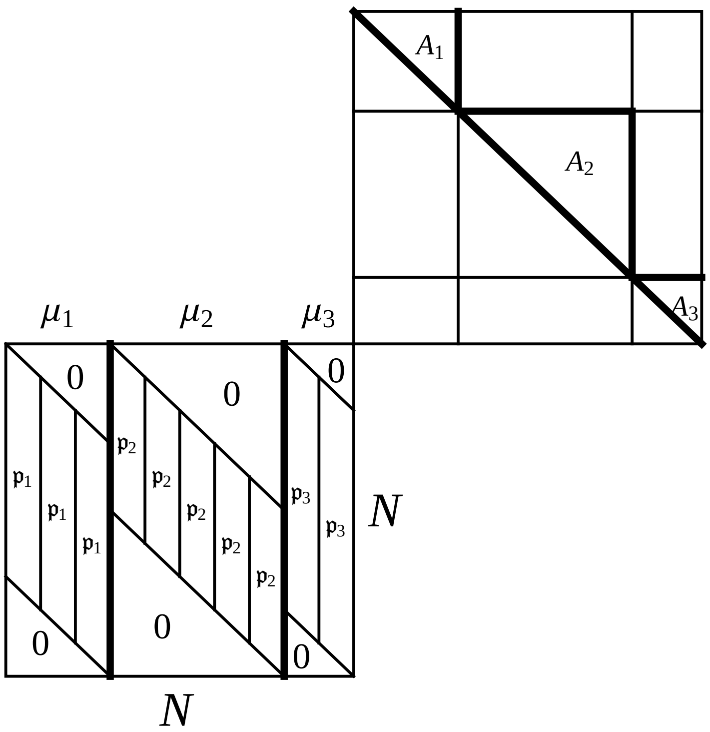
$$
\underbrace{[\mathrm{D}_{\mu_1-1}(\mathfrak{p}_1) | \dots | \mathrm{D}_{\mu_k-1}(\mathfrak{p}_k)]}_{B} \cdot \underbrace{\mathrm{diag}\{A_1, \dots A_k\}}_{A} \cdot \underline{a} = \underbrace{(1, 0, \dots 0)^\top}_{\underline{e}} \tag{10}
$$

egyenlet megoldását keressük. Látható, hogy bármilyen $\underline{e} \in \mathbb{R}^N$ vektort is írhatnánk a jobb oldalra, vagyis bármilyen $q(x) \in \mathbb{R}_{N-1}[x]$ polinom is lehetett volna $\frac{1}{p(x)}$ számlálójában.

Vizsgáljuk még (10) képletet. $\mathrm{D}_{\mu_i-1}(\mathfrak{p}_i) \in \mathbb{R}^{N \times \mu_i}$, mert $\mathfrak{p}_i$ fokszáma $N - \mu_i$ (elevenítsük fel (2) képletet). Ugyanakkor A_i is $\mu_i \times \nu_i$-es, ezért szorozhatóak. Az A mátrix blokk-felsőháromszög alakú, csupa egyessel a főátlóban, míg B egymás mellé rakott sáv szerkezetű mátrixokból áll, $B \in \mathbb{R}^{N \times (\mu_1 + \dots + \mu_k)} = \mathbb{R}^{N \times N}$

4. A bizonyítás

Eddig a (7) feladatot ekvivalensen megfogalmaztuk $B \cdot A \cdot \underline{a} = \underline{e}$ alakban a fent megkonstruált mátrixok és vektorok segítéségvel. A-ról triviálisan láttuk, hogy invertálható, nem maradt más hátra, mint hogy belássuk, $\det(B) \neq 0 \Leftrightarrow \mathrm{rank}(B) = N$.



1. ábra. A $B \cdot A$ szorzat struktúrája

4.1. Tétel. *Ha egy $p(x) \in \mathbb{R}_N[x]$ polinomot a (3) alakban faktorizálunk, majd a faktorokból megkonstruáljuk B mátrixot, akkor B determinánsa nem 0.*
Következésképp (7) egyenletnek $\exists!\underline{a} \in \mathbb{R}^N$ megoldása bármilyen $q(x) \in \mathbb{R}_{N-1}[x]$ számláló is álljon a bal oldalon.

Allításunk ekvivalens azzal, hogy

$$\left\langle \mathfrak{p}_1(x), x \cdot \mathfrak{p}_1(x), \dots x^{\mu_1 - 1}\mathfrak{p}_1(x), \dots \mathfrak{p}_k(x), , \dots x^{\mu_k - 1}\mathfrak{p}_k(x) \right\rangle = \mathbb{R}_{N-1}[x].$$

Építsük fel B mátrixot lépésenként és lássuk, hogy a rang teljes marad.

$$B_2 := [\mathrm{D}_{\mu_1 - 1}(P_2) | \, \mathrm{D}_{\mu_2 - 1}(P_1)]$$

Először lássuk, hogy B_2 teljes rangú ($k \geq 2$ miatt B_2 értelmes). $\mathrm{rank}(B_2) \geq \mu_1$, ezt $\mathrm{D}_{\mu_1 - 1}(P_2)$ biztosítja.

Teljes rang azt jelenti, hogy egyik oszlopa sem állítható elő a többi lineáris kombinációjával. Először nézzük meg, hogy a μ_1+1-edik oszlop előállítható-e a megelőzőekkel.

$$\left\langle P_2(x), x\cdot P_2(x), \dots x^{\mu_1-1}P_2(x)\right\rangle = \left\langle 1, x, \dots x^{\mu_1-1}\right\rangle P_2(x) = \tag{11}$$

$$\mathbb{R}_{\mu_1-1}[x]\cdot P_2(x) \tag{12}$$

de

$$P_1(x) \notin \mathbb{R}_{\mu_1-1}[x]\cdot P_2(x) \tag{13}$$

mert P_1 nem tartalmazza P_2-t mint faktor és $P_1(x) \notin \mathbb{R}_{\mu_1-1}[x]$. Tehát $\mathrm{rank}(B_2) \geq \mu_1+1$.

A többi oszlop előállíthatóságát is megvizsgáljuk. Vegyük az oszlopokat $l = 1\dots\mu_2$-ig. Tegyük fel indirekt, hogy $x^{l-1}\cdot P_1(x) = q_{\mu_1-1}(x)\cdot P_2(x) + q_{l-2}(x)\cdot P_1(x)$ valamilyen $q_{\mu_1-1}\in\mathbb{R}_{\mu_1-1}[x]$ ill. $q_{l-2}\in\mathbb{R}_{l-2}[x]$ polinomokra (ez pont az $1\dots(\mu_1+l-1)$ oszlopok lineár kombinációja).

$$x^{l-1}\cdot P_1(x) = q_{\mu_1-1}(x)\cdot P_2(x) + q_{l-2}(x)\cdot P_1(x) \tag{14}$$

$$\Updownarrow$$

$$\underbrace{\left(x^{l-1} - q_{l-2}(x)\right)}_{\in\mathbb{R}_{l-1}[x]\subseteq\mathbb{R}_{\mu_2-1}[x]}\cdot P_1(x) = q_{\mu_1-1}(x)\cdot P_2(x) \tag{15}$$

A két oldal nem lehet egyenlő, mert $P_1(x)$ és $P_2(x)$-ben *nincsenek közös faktorok* és egy alacsonyabb fokú nem tud egy egyel magasabb fokút generálni.

Most úgy bővítjük B_2-t, hogy

$$B_{n+1} := \left[{}_{N-1}\mathrm{D}(P_{n+1})\cdot B_n \middle| \mathrm{D}_{\mu_{n+1}-1}\left(\prod_{i=1}^{n} P_i(x)\right)\right] \tag{16}$$

Magyarán mindent beszorzunk az új polinommal és mögé hozzávesszük az eddigiek szorzatát (az új nélkül). Ezzel pont B-t állítjuk elő:

$$B_n = \left[\mathrm{D}_{\mu_1-1}\left(\prod_{i=2\dots n} P_i(x)\right) \middle| \dots \mathrm{D}_{\mu_l-1}\left(\prod_{\substack{i=1\dots n\\ i\neq l}} P_i(x)\right) \dots \middle| \mathrm{D}_{\mu_n-1}\left(\prod_{i=1\dots n-1} P_i(x)\right)\right]$$

ahol $n = 2\dots k$ és $B_k = B$.

Lássuk, hogy a (16) konstrukcióval minden lépésben μ_{n+1}-el nő a rang. Indirekt tegyük fel, hogy nem nő valahol a rang, vagyis az $n+1$-edik blokk l-edik oszlopa előáll ${}_{N-1}\mathrm{D}(P_{n+1})\cdot B_n$ oszlopainak és az $n+1$-edik blokk $1 \ldots (l-1)$ oszlopainak lineár kombinációjaként.

Vagyis

$$x^{l-1} \cdot \prod_{i=1\ldots n} P_i(x) =$$

$$\begin{aligned}
&q_{\mu_1-1}(x) \cdot \prod_{i=2\ldots n+1} P_i(x) + \ldots + \\
&q_{\mu_l-1}(x) \cdot \prod_{\substack{i=1\ldots n+1 \\ i\neq l}} P_i(x) + \ldots + \\
&q_{\mu_n-1}(x) \cdot \prod_{\substack{i=1\ldots n+1 \\ i\neq n}} P_i(x) + \\
&q_{l-2}(x) \cdot \prod_{i=1\ldots n} P_i(x)
\end{aligned}$$

valamilyen $q_\bullet \in \mathbb{R}_\bullet[x]$ polinomokra. Vigyük át a jobb oldakról az utolsó tagot.

$$\overbrace{\left(x^{l-1} - q_{l-2}(x)\right)}^{\in\mathbb{R}_{l-1}[x]\subseteq\mathbb{R}_{\mu_{n+1}-1}[x]} \cdot \prod_{i=1\ldots n} P_i(x) =$$

$$\begin{aligned}
&q_{\mu_1-1}(x) \cdot \prod_{i=2\ldots n+1} P_i(x) + \ldots + \\
&q_{\mu_l-1}(x) \cdot \prod_{\substack{i=1\ldots n+1 \\ i\neq l}} P_i(x) + \ldots + \\
&q_{\mu_n-1}(x) \cdot \prod_{\substack{i=1\ldots n+1 \\ i\neq n}} P_i(x) =
\end{aligned}$$

$$= P_{n+1}(x) \cdot \Bigg(\qquad \Bigg)$$

A nagy zárójel belsejében nem jelenik meg a teljes $\prod_{i=1\ldots n} P_i(x)$ csak hiányos faktorokkal és az egyes tagokban amelyik faktor hiányzik, annál egyel alacsonyabb rendű q polinommal van szorozva. Továbbra is kihasználjuk, hogy P_i-k páronként relatív prímek, így ez az egyenlőség nem jöhet létre.

Egyben azt is beláttuk, hogy bármilyen $P_i(x) \in \mathbb{R}_{\mu_i}[x]$, $i = 1\ldots k$ páronként relatív prím polinomokra, ahol $\mu_i > 0$ és $\sum \mu_i = N$:

$$\left\langle \mathbb{R}_{\mu_l-1}[x] \cdot \prod_{i\neq l} P_i(x) \right\rangle_{l=1\ldots k} = \mathbb{R}_{N-1}[x] \qquad \square$$