Téglalap alakú mátrixok szinguláris felbontása.

Legyen adva egy $p \times q$ méretű $\mathbf{A}$ mátrix. Meg akarjuk adni ennek az $\mathbf{A}$ mátrixnak a legegyszerűbb előállítását. Ezt nevezik az $\mathbf{A}$ mátrix szinguláris felbontásának. Négyzetes, azaz $p \times p$ méretű $\mathbf{A}$ mátrixok esetén létezik egy $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}$ alakú előállítás, ahol $\boldsymbol{\Lambda}$ $p \times p$ méretű diagonális mátrix, az átlóban $\lambda_1 \geq \lambda_1 \geq \cdots \geq \lambda_p \geq 0$ nem negatív elemekkel, (azaz $\boldsymbol{\Lambda}$ pozitív definit mátrix), és $\mathbf{U}$ és $\mathbf{V}$ ortogonális mátrixok az R^p térben. Ennek az eredménynek keressük az általánosítását téglalap alakú mártrixokra.

Ezt az eredményt három ekvivalens alakban fogom megfogalmazni. Ezek tárgyalása előtt megjegyzem, hogy amennyiben $\mathbf{A}$ $p \times q$ méretű mátrix, akkor az $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}\mathbf{A}^T$ és $\mathbf{A}^T\mathbf{A}$ mátrixok s rangjára igaz, hogy $s = \operatorname{rang}(\mathbf{A}) = \operatorname{rang}(\mathbf{A}^T) = \operatorname{rang}(\mathbf{A}\mathbf{A}^T) = \operatorname{rang}(\mathbf{A}^T\mathbf{A}) \leq \min(p,q)$. Ahhoz, hogy a

$$\operatorname{rang}(\mathbf{A}) = \operatorname{rang}(\mathbf{A}\mathbf{A}^T)$$

relációt belássuk, megmutatjuk, hogy $\mathbf{u}^T\mathbf{A}\mathbf{A}^T = \mathbf{0}$ esetén $\mathbf{u}^T\mathbf{A} = \mathbf{0}$, mert ekkor $(\mathbf{u}^T\mathbf{A}, \mathbf{u}^T\mathbf{A}) = (\mathbf{u}^T\mathbf{A}\mathbf{A}^T, \mathbf{u}) = 0$.

Az $\mathbf{A}$ mátrix szinguláris felbontásának első megfogalmazása a következő. Legyen a $p \times q$ méretű $\mathbf{A}$ mátrix rangja s. Ekkor létezik s darab p hosszúságú $\mathbf{u}_1, \dots, \mathbf{u}_s$ ortonormális (oszlop)vektor, s darab q hosszúságú $\mathbf{v}_1, \dots, \mathbf{v}_s$ ortonormális (oszlop)vektor és $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s > 0$ számok úgy, hogy az $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_s)$ $p \times s$ méretű, a $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_s)$ $q \times s$ méretű valamint a $\boldsymbol{\Lambda}$ $s \times s$ méretű diagonális mátrix $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s > 0$ elemekkel az átlóban teljesítik az $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T$ azonosságot.

Az $\mathbf{A}$ mátrix szinguláris felbontásának második megfogalmazásában az első megfogalmazásban szereplő $\mathbf{U}$ és $\mathbf{V}$ mátrixokat kiegészítem $p \times p$ és $q \times q$ méretű ortogonális mátrixokká, a $\boldsymbol{\Lambda}$ mátrixot pedig nulla elemű koordináták hozzáadásával egy $p \times q$ méretű mátrixszá. Az így kapott mátrixokat továbbra is az $\mathbf{U}$, $\mathbf{V}$ és $\boldsymbol{\Lambda}$ betűkkel fogom jelölni. Azt állítom, hogy ezek a mátrixok is teljesítik a szinguláris felbontás első megfogalmazásában felírt azonosságot.

Részletesebben megfogalmazva egészítsük ki az $\mathbf{u}_1, \dots, \mathbf{u}_s$ vektorrendszert egy $\mathbf{u}_1, \dots, \mathbf{u}_p$ ortonormált bázissá az R^p térben. Hasonlóan egészítsük ki a $\mathbf{v}_1, \dots, \mathbf{v}_s$ vektorrendszert egy $\mathbf{v}_1, \dots, \mathbf{v}_q$ ortonormált bázissá az R^q térben. Legyen $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_p)$ és $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_q)$. Végül legyen $\boldsymbol{\Lambda}$ az a $p \times q$ méretű $\boldsymbol{\Lambda} = (a_{i,j})$, $1 \leq i \leq p$, $1 \leq j \leq q$, mátrix, amelyre $a_{i,i} = \lambda_i$, ha $1 \leq i \leq s$, és $a_{i,j} = 0$, ha $i \neq j$ vagy $i > s$. Ezzel a választással teljesül az $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T$ azonosság.

A szinguláris felbontás e két megfogalmazásának az ekvivalenciája egyszerűen látható. Azt kell észrevenni, hogy a második megfogalmazásban szereplő Λ mátrix alakjából következik, hogy az ott szereplő $\mathbf{U\Lambda V}^T$ szorzat értéke nem függ attól, hogy hogyan választottuk meg az $\mathbf{u}_{s+1},\dots,\mathbf{u}_p$ és $\mathbf{v}_{s+1},\dots,\mathbf{v}_q$ vektorokat. Innen látható, hogy ha adott egy a szinguláris felbontás második jellemzését teljesítő rendszer, akkor elhagyva az $\mathbf{U}$ mátrix utolsó $p-s$ oszlopvektorát, a $\mathbf{V}$ mátrix utolsó $q-s$ oszlopvektorát, és a $\mathbf{\Lambda}$ mátrixot megszorítjuk az első s sorára és oszlopára, akkor olyan $\mathbf{U}$, $\mathbf{V}$ és $\mathbf{\Lambda}$ mátrixokat kapunk, amelyek teljesítik a szinguláris felbontás első jellemzését.

Megfordítva, ha egy $\mathbf{U}$, $\mathbf{V}$, $\mathbf{\Lambda}$ rendszer teljesíti a szinguláris felbontás első jellemzését, akkor kiegészítve az $\mathbf{U}$ és $\mathbf{V}$ mátrixokat ortogonális mátrixokká olyan módon, ahogy azt leírtam, és a $\mathbf{\Lambda}$ mátrixot szintén a korábban leírt módon egészítve ki egy $p\times q$ méretű mátrixszá egy olyan $\mathbf{U}$, $\mathbf{V}$, $\mathbf{\Lambda}$ rendszert kapunk, amely teljesíti a szinguláris felbontás második jellemzését.

A szinguláris felbontás létezéséről szóló eredmény harmadik megfogalmazása az $\mathbf{x}\in R^p$ és $\mathbf{y}\in R^q$ vektorpárok $\alpha(\mathbf{x},\mathbf{y})$ bilineáris függgvényeinek egyszerű megadásáról szól. Egy $\alpha(\mathbf{x},\mathbf{y})$, $\mathbf{x}\in R^p$, $\mathbf{y}\in R^q$, függvény bilineáris, ha tetszőleges a és b számokra és $\mathbf{x_1}\in R^p$, $\mathbf{x}_2\in R^p$, $\mathbf{x}\in R^p$, $\mathbf{y_1}\in R^q$, $\mathbf{y}_2\in R^q$, $\mathbf{y}\in R^q$ vektorokra az $\alpha(a\mathbf{x}_1+b\mathbf{x}_2,\mathbf{y})=a\alpha(\mathbf{x}_1,\mathbf{y})+b(\alpha\mathbf{x}_2,\mathbf{y})$ és $\alpha(a\mathbf{x},a\mathbf{y}_1+b\mathbf{y}_2)=a\alpha(\mathbf{x},\mathbf{y}_1)+b\alpha(\mathbf{x},\mathbf{y}_2)$ azonosságok teljesülnek.

Egy $\mathbf{A}$ $p\times q$ méretű mátrix meghatároz egy $\alpha(\mathbf{x},\mathbf{y})$, $\mathbf{x}\in R^p$, $\mathbf{y}\in R^q$, bilineáris függvényt a következő módon. Ha $\mathbf{x}\in R^p$, $\mathbf{y}\in R^q$, akkor legyen $\alpha(\mathbf{x},\mathbf{y})=\mathbf{x}^T\mathbf{A}\mathbf{y}$. Másrészt az is igaz, hogy minden $\alpha(\mathbf{x},\mathbf{y})$, $\mathbf{x}\in R^p$, $\mathbf{y}\in R^q$, bilineáris függvény megadható ilyen alakban.

Ezt megmutatandó definiáljuk az $\mathbf{e}_i\in R^p$, $1\le i\le p$, vektorokat, ahol $\mathbf{e}_i$ az a vektor az R^p térben, amelynek az i-ik koordinátája 1 a többi koordinátája 0. Hasonlóan, vezessük be a $\mathbf{f}_j\in R^q$, $1\le j\le q$, vektorokat, ahol az $\mathbf{f}_j$ vektor j-ik koordinátája 1 a többi koordinátája nulla. Ekkor az $\mathbf{e}_1,\dots,\mathbf{e}_p$ vektorok bázist alkotnak az R^p térben, az $\mathbf{f}_1,\dots,\mathbf{f}_q$ vektorok bázist alkotnak az R^q térben. Továbbá, egy $\alpha(\mathbf{x},\mathbf{y})$, $\mathbf{x}\in R^p$, $\mathbf{y}\in R^q$, bilineáris függvény jellemezhető a következő módon.

Tekintsünk egy $\alpha(\mathbf{e}_i,\mathbf{f}_j)=A(i,j)$, $1\le i\le p$, $1\le j\le q$, függvényt tetszőleges $A(i,j)$ konstansokkal. Létezik olyan $\alpha(\mathbf{x},\mathbf{y})$, $\mathbf{x}\in R^p$, $\mathbf{y}\in R^q$, bilineáris függvény, amelyre $\alpha(\mathbf{e}_i,\mathbf{f}_j)=A(i,j)$, $1\le i\le p$, $1\le j\le q$, és az megadható a következő képlet segítségével: $\alpha(\mathbf{x},\mathbf{y})=\mathbf{x}^T\mathbf{A}\mathbf{y}$, ahol az $\mathbf{A}$ mátrixot az $\mathbf{A}=(A(i,j))$, $1\le i\le p$, $1\le j\le q$, képlet határozza meg. Továbbá minden $\alpha(\mathbf{x},\mathbf{y})$, $\mathbf{x}\in R^p$, $\mathbf{y}\in R^q$, bilineáris függvény megadható ilyen alakban.

Tekintsünk egy $\alpha(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}$ bilineáris függvényt, és vegyünk egy $\mathbf{x}_1 \in R^p,\dots,\mathbf{x}_p \in R^p$ bázist az R^p térben, és egy $\mathbf{y}_1 \in R^q,\dots,\mathbf{y}_q \in R^q$ bázist az R^q térben. Ekkor az $\alpha(\mathbf{x}_i,\mathbf{y}_j)$, $1 \le i \le p$, $1 \le j \le q$, mennyiségek azaz az $\alpha(\mathbf{x},\mathbf{y})$ függvény értékei az $\mathbf{x}_i \in R^p$, $1 \le i \le p$, $\mathbf{y}_j \in R^q$, $1 \le j \le q$, bázispáron meghatározzák az $\alpha(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}$ bilineáris függvényt. A szinguláris felbontás harmadik megfogalmazásában olyan bázispárt adunk meg, amelyben a bilineáris függvény ilyen módon való megadása egyszerű szerkezetű.

A szinguláris felbontás harmadik megfogalmazása a következőt állítja. Legyen adva egy $\alpha(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}$, $\mathbf{x} \in R^p$, $\mathbf{y} \in R^q$, bilineáris függvény, ahol az $\mathbf{A}$ mátrix rangja s. Ekkor létezik olyan $\mathbf{u}_1 \in R^p,\dots,\mathbf{u}_p \in R^p$, ortonormált bázis az R^p térben, $\mathbf{v}_1 \in R^q,\dots,\mathbf{v}_q \in R^q$, ortonormált bázis az R^q térben, és léteznek olyan $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_s > 0$ számok, amelyekre $\alpha(\mathbf{u}_i,\mathbf{v}_j) = 0$, ha $i \ne j$, $\alpha(\mathbf{u}_i,\mathbf{v}_i) = 0$, ha $i > s$, és $\alpha(\mathbf{u}_i,\mathbf{v}_i) = \lambda_i$, ha $1 \le i \le s$.

Lássuk be, hogy a szinguláris felbontás második megfogalmazásában szereplő állításból következik a szinguláris felbontás harmadik megfogalmazásában megfogalmazott állítás. Legyen adva egy $\alpha(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}$ bilineáris függvény, és írjuk az $\mathbf{A}$ mátrixot $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$ alakban. Legyenek az $\mathbf{U}$ mátrix oszlopvektorai az $\mathbf{u}_1,\dots,\mathbf{u}_p$ vektorok, és a $\mathbf{V}$ mátrix oszlopvektorai a $\mathbf{v}_1,\dots,\mathbf{v}_q$ vektorok. Azt állítom, hogy ezekkel a $\mathbf{u}_1,\dots,\mathbf{u}_p$ és $\mathbf{v}_1,\dots,\mathbf{v}_q$ vektorokkal, valamint a $\mathbf{\Lambda}$ mátrix átlójában szereplő λ_j számokkal érvényes az $\alpha(\mathbf{x},\mathbf{y})$ bilineáris függvénynek a szinguláris felbontás harmadik megfogalmazásában szereplő jellemzése.

Ennek megmutatása érdekében vegyük észre, hogy $\mathbf{u}_i^T\mathbf{U} = \mathbf{e}_i^T$ $1 \le i \le p$, és $\mathbf{V}^T\mathbf{v_j} = \mathbf{f}_j$, $1 \le j \le q$. Ezért $\alpha(\mathbf{u}_i,\mathbf{v}_j) = \mathbf{u}_i^T\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T\mathbf{v}_j = \mathbf{e}_i^T\mathbf{\Lambda}\mathbf{f}_j = \delta_{i,j}\lambda_j$, ahol $\delta_{i,j} = 1$, ha $i = j$, $\delta_{i,j} = 0$, ha $i \ne j$. Innen következik a szinguláris felbontás harmadik megfogalmazásában megfogalmazott állítás.

Megmutatjuk azt is, hogy ha a szinguláris felbontás harmadik megfogalmazásában szereplő állítás igaz, akkor a szinguláris felbontás második megfogalmazásában megfogalmazott állítás is igaz.

Legyen adva egy $\mathbf{A}$ $p \times q$ méretú mátrix. Be akarjuk látni, hogy érvényes rá a szinguláris felbontás második megfogalmazásában szereplő állítás. Ennek érdekében tekintsük az e mátrix által meghatározott $\alpha(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}$ bilineáris függvényt. A szinguláris felbontás harmadik megfogalmazása szerint létezik olyan $\mathbf{u}_1 \in R^p,\dots,\mathbf{u}_p \in R^p$, ortonormált bázis az R^p térben, $\mathbf{v}_1 \in R^q,\dots,\mathbf{v}_q \in R^q$, ortonormált bázis az R^q térben, és léteznek olyan $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_s > 0$ számok, amelyekre $\alpha(\mathbf{u}_i,\mathbf{v}_j) = 0$, ha $i \ne j$, $\alpha(\mathbf{u}_i,\mathbf{v}_i) = 0$, ha $i > s$, és $\alpha(\mathbf{u}_i,\mathbf{v}_i) = \lambda_i$, ha $1 \le i \le s$. Definiáljuk az

$\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_p)$ és $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_q)$ mátrixokat, valamint a $\mathbf{\Lambda} = (a_{i,j})$, $1 \le i \le p$, $1 \le j \le q$, mátrixot ahol $a_{i,j} = 0$, ha $i \ne j$, $a_{i,i} = 0$, ha $i > s$, $a_{i,i} = \lambda_i$, ha $1 \le i \le s$. Definiáljuk továbbá a $\mathbf{B} = \mathbf{U}^T \mathbf{\Lambda} \mathbf{V}$ mátrixot. Ekkor $\mathbf{u}_i^T \mathbf{B} \mathbf{v}_j = \alpha(\mathbf{u}_i, \mathbf{v}_j) = \mathbf{u}_i^T \mathbf{A} \mathbf{v}_j$ minden $1 \le i \le p$, $1 \le j \le q$ indexre. Innen következik, hogy $\mathbf{e}_i^T \mathbf{B} \mathbf{f}_j = \mathbf{e}_i^T \mathbf{A} \mathbf{f}_j$ minden $1 \le i \le p$, $1 \le j \le q$ indexre. Ezért $\mathbf{A} = \mathbf{B} = \mathbf{U}^T \mathbf{\Lambda} \mathbf{V}$. Ez azt jelenti, hogy a szinguláris felbontás harmadik megfogalmazásában szereplő jellemzéséből következik a szinguláris felbontás második megfogalmazásában szereplő jellemzése.

A szinguláris felbontásról szóló állítás harmadik verzióját fogom bizonyítani.

Legyen adva egy $p \times q$ méretű $\mathbf{A}$ mátrix, amelynek rangja s, és definiáljuk segítségével az $\alpha(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y}$, $\mathbf{x} \in R^p$, $\mathbf{y} \in R^q$, bilineáris függvényt. Megadom, hogy hogyan definiáljuk az eredmény megfogalmazásában szereplő $\mathbf{u}_1 \in R^p, \dots, \mathbf{u}_p \in R^p$, illetve $\mathbf{v}_1 \in R^q, \dots, \mathbf{v}_q \in R^q$, vektorokat, és $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_s > 0$ számokat.

Az $\mathbf{A}^T \mathbf{A}$ mátrix szimmetrikus és pozitív definit az R^q térben, rangja s. Ezért léteznek $\mathbf{v}_1 \in R^q, \dots, \mathbf{v}_q \in R^q$ ortonormált sajátvektorai, $\mu_1 \ge \mu_2 \ge \cdots \mu_q \ge 0$ sajátértékekkel, és $\mu_j = 0$, ha $q \ge j > s$. Továbbá , ha $j > s$, akkor nemcsak $\mathbf{A}^T \mathbf{A} \mathbf{v}_j = \mathbf{0}$, hanem $\mathbf{A} \mathbf{v}_j = \mathbf{0}$. Ezt hasonlóan lehet belátni, mint ahogy megmutattam az ismertetés elején, hogy $\mathbf{u}^T \mathbf{A} = \mathbf{0}$, ha $\mathbf{u}^T \mathbf{A} \mathbf{A}^T = \mathbf{0}$.

A $\mathbf{v}_1, \dots, \mathbf{v}_q$ vektorokat már definiáltuk. Legyen $\lambda_j = \sqrt{\mu_j}$, $1 \le j \le s$. Az $\mathbf{u}_j$ vektorokat az első lépésben csak $1 \le j \le s$ indexekre definiáljuk. Legyen $\mathbf{u_j} = \frac{1}{\lambda_j} \mathbf{A} \mathbf{v}_j$, ha $1 \le j \le s$. Azt állítom, hogy az $\mathbf{u}_j$, $1 \le j \le s$, vektorok ortonomált rendszert alkotnak, és $\mathbf{A} \mathbf{A}^T \mathbf{u}_j = \mu_j \mathbf{u}_j$, ha $1 \le j \le s$.

Valóban, $(\mathbf{u}_i, \mathbf{u}_j) = \mathbf{u}_i^T \mathbf{u}_j = \frac{1}{\lambda_i \lambda_j} \mathbf{v}_i^T \mathbf{A}^T \mathbf{A} \mathbf{v}_j = \frac{1}{\lambda_i \lambda_j} \mathbf{v}_i^T \mu_j \mathbf{v}_j = \frac{\lambda_j^2}{\lambda_i \lambda_j} \mathbf{v}_i^T \mathbf{v}_j = \delta_{i,j}$, ha $1 \le i, j \le s$, ahol $\delta_{i,j} = 1$, ha $i = j$, $\delta_j = 0$, ha $i \ne j$. Ebben a számolásban felhasználtuk, hogy $\mathbf{A}^T \mathbf{A} \mathbf{v}_j = \mu_j \mathbf{v}_j$, és $\mathbf{v}_i^T \mathbf{v}_j = \delta_{i,j}$.

Másrészt $\mathbf{A} \mathbf{A}^T \mathbf{u}_j = \frac{1}{\lambda_j} \mathbf{A} \mathbf{A}^T \mathbf{A} \mathbf{v}_j = \frac{1}{\lambda_j} \mathbf{A} (\mathbf{A}^T \mathbf{A} \mathbf{v}_j) = \frac{\mu_j}{\lambda_j} \mathbf{A} \mathbf{v}_j = \mu_j \mathbf{u}_j$, ha $1 \le j \le s$. Ez azt jelenti, hogy az $\mathbf{u}_i$, $1 \le i \le s$, vektorok ortonormáltak és sajátvektorai $\mu_i > 0$ sajátértékkel az $\mathbf{A} \mathbf{A}^T$ szimmetrikus mátrixnak. Mivel az $\mathbf{A} \mathbf{A}^T$ mátrix rangja s, ezért a $(\mathbf{u}_1, \dots, \mathbf{u}_s)$ vektorok rendszerét kiegészíthetjük egy az $\mathbf{A} \mathbf{A}^T$ mátrix sajátvektoraiból álló $(\mathbf{u}_1, \dots, \mathbf{u}_p)$ ortonormált bázissá, ahol $\mathbf{A} \mathbf{A}^T \mathbf{u}_j = \mathbf{0}$, ha $s < j \le p$. Ezzel definiáltuk az $\mathbf{u}_1, \dots, \mathbf{u}_p$ vektorokat is. Azt álllítom, hogy ez a rendszer teljesíti a szinguláris felbontásról szóló állítás harmadik verzióját.

Azt kell még megmutatni, hogy $\alpha(\mathbf{u}_i, \mathbf{v}_j) = \mathbf{u}_i^T \mathbf{A} \mathbf{v}_j = \delta_{i,j} \lambda_j$ minden $1 \le$

$i \le p$, $1 \le j \le q$, indexre. Ha $1 \le j \le s$, akkor $\mathbf{u}_i^T \mathbf{A} \mathbf{v}_j = \mathbf{u}_i^T (\lambda_j \mathbf{u}_j) = \delta_{i,j}\lambda_j$. Ha $s < j \le q$, akkor $\mathbf{A}\mathbf{v}_j = \mathbf{0}$, ezért $\mathbf{u}_i^T \mathbf{A}\mathbf{v}_j = 0$. Másrészt ekkor $\lambda_j = 0$, ezért $\delta_{i,j}\lambda_j = 0$. Ezzel a kívánt relációt és a szinguláris felbontásról szóló állítás harmadik verzióját beláttuk.

Ebben az írásban valós értékű mátrixok, azaz olyan mátrixok jellemzését adtuk meg, amelyek az R^p tér R^q térbe való lineáris leképezéseinek felelnek meg. Hasonlóan lehet megadni (és bizonyítani) komplex értékű mátrixok jellemzését, amelyek a Z^p p-dimenziós komplex térnek a Z^q q-dimenziós komplex térbe való lineáris leképezéseinek felelnek meg. A fő különbség az, hogy ekkor egy $\mathbf{A}$ mátrix $\mathbf{A}^T$ transzponáltja helyett annak $\mathbf{A}^*$ konjugáltját, illetve egy $\mathbf{x}$ vektor $\mathbf{x}^T$ transzponáltja helyett annak $\mathbf{x}^*$ konjugáltját vesszük. Egy másik különbség az, hogy a szinguláris felbontás harmadik verziójának megfogalmazásában bilineáris függvények helyett komplex bilineáris függvények jellemzését adjuk meg, azaz olyan $\alpha(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \in Z^p$, $\mathbf{y} \in Z^q$, függvények jellemzését, amelyekre az $\alpha(a\mathbf{x}_1 + b\mathbf{x}_2, \mathbf{y}) = a\alpha(\mathbf{x}_1, \mathbf{y}) + b(\alpha\mathbf{x}_2, \mathbf{y})$ és $\alpha(a\mathbf{x}, a\mathbf{y}_1 + b\mathbf{y}_2) = \bar{a}\alpha(\mathbf{x}, \mathbf{y}_1) + \bar{b}\alpha(\mathbf{x}, \mathbf{y}_2)$ azonosságok teljesülnek tetszőleges $a \in Z$ és $b \in Z$ számokra és $\mathbf{x}_1 \in Z^p$, $\mathbf{x}_2 \in Z^p$, $\mathbf{x} \in Z^p$, $\mathbf{y}_1 \in Z^q$, $\mathbf{y}_2 \in Z^q$, $\mathbf{y} \in Z^q$ vektorokra. Az irodalomban komplex bilineáris függvények helyett gyakran sesquilineáris (másfél lineáris) függvényekről beszélnek.

Ha komplex értékű mátrixokat vizsgálunk, akkor érdemesebb sor és nem oszlopvektorokkal dolgozni. A továbbiakban ezt fogjuk tenni. Ilyen jelöléssel az $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}\mathbf{y}^*$ skalárszorat, $\mathbf{x} \in Z^n$, $\mathbf{y} \in Z^n$, példa komplex bilineáris függvényre. Egy általános $\alpha(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \in Z^p$, $\mathbf{y} \in Z^q$ komplex bilineáris függvény $\alpha(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}\mathbf{A}, \mathbf{y} \rangle = \mathbf{x}\mathbf{A}\mathbf{y}^*$ alakban adható meg, ahol $\mathbf{A}$ $p \times q$ méretű mátrix, amelynek elemei komplex számok.

Megfogalmazom a szinguláris felbontásról szóló állítás harmadik verzióját abban az esetben, ha komplex értékű mátrixokkal dolgozunk. Ezután röviden ismertetem a bizonyítás gondolatát.

Először teszek egy apró megjegyzést. Érdemes észrevenni, hogy egy $\mathbf{x}$ sorvektor akkor és csak akkor sajátvektora egy $\mathbf{A}$ szimmetrikus mátrixnak egy (valós) λ sajátértékkel, ha a konjugáltja az $\mathbf{x}^*$ oszlopvektor sajátvektora ugyanannak az $\mathbf{A}$ szimmetrikus mátrixnak ugyanazzal a λ sajátértékkel, azaz $\mathbf{x}\mathbf{A} = \lambda\mathbf{x}$ akkor és csak akkor, ha $\mathbf{A}\mathbf{x}^* = \lambda\mathbf{x}^*$.

A mátrixok szinguláris felbontásáról szóló állítás harmadik verziója a következő módon fogalmazható meg.

Legyen adva egy $\alpha(\mathbf{x}, \mathbf{y}) = \mathbf{x}\mathbf{A}\mathbf{y}^*$ komplex bilineáris függvény, ahol $\mathbf{x} \in Z^p$, $\mathbf{y} \in Z^q$, sorvektorok, és $\mathbf{A}$ olyan $p \times q$ méretű komplex számokat tartalmazó mátrix, amelynek rangja s . Ekkor léteznek olyan $\mathbf{u}_1 \in Z^p, \dots, \mathbf{u}_p \in Z^p$ a Z^p térben ortonormált bázist alkotó (sor)vektorok, $\mathbf{v}_1 \in Z^q, \dots, \mathbf{v}_q \in Z^q$ a Z^q térben ortonormált bázist alkotó (sor)vektorok és $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_s > 0$ számok, amelyekre $\alpha(\mathbf{u}_i, \mathbf{v}_j) = \mathbf{u}_i \mathbf{A} \mathbf{v}_j^* = \delta_{i,j}\lambda_j$ minden $1 \le i \le p$ és $1 \le j \le q$ indexre, ahol $\lambda_j = 0$, ha $s < j \le q$.

A fenti állításban szereplő vektorokat és konstansokat a következő módon definiálhatjuk. Legyen $\mathbf{v}_j$ az $\mathbf{A}^*\mathbf{A}$ mátrix j-edik sajátvektora (sorvektora) μ_j sajátértékkel, és válasszuk e vektorokat úgy, hogy ezek a $\mathbf{v}_j$, $1 \le j \le q$, vektorok ortonormált bázist alkotnak a Z^q térben. Legyen $\lambda_j = \sqrt{\mu_j}$, $1 \le j \le s$. Legyen $\mathbf{u}_j = \frac{1}{\lambda_j}\mathbf{v}_j \mathbf{A}^*$, ha $1 \le j \le s$. Ekkor az $\mathbf{u}_j$, $1 \le j \le s$, vektorok ortonormáltak, és sajátvektorai a $\mathbf{A}\mathbf{A}^*$ szimmetrikus mátrixnak. Ezeket az $\mathbf{u}_j$, $1 \le j \le s$, vektorokat kiterjesztjük az $\mathbf{A}\mathbf{A}^*$ mátrix sajátvektoraiból álló $\mathbf{u}_j$, $1 \le j \le p$, teljes ortonormált rendszerré a Z^p térben. Így választjuk a $\mathbf{u}_j$, $1 \le j \le p$, vektorokat. Ezzel a választással érvényes a szinguláris felbontásról szóló állítás harmadik verziója komplex értékű mátrixokra.

Visszatérek a skalárértékű bilineáris függvények tulajdonságainak a tárgyalásához. Megfogalmazok néhány számunkra érdekes állítást, és elmagyarázom, hogy azok hogyan kapcsolódnak a már bizonyított eredményekhez. A bizonyítások részleteit nem dolgozom ki. Azt az érdeklődő olvasó maga is megteheti. Az állítások alkalmas megfelelője igaz komplex értékű mátrixokra is.

Ha adva van egy $\alpha(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \in R^p$, $\mathbf{y} \in R^q$, bilineáris függvény, akkor annak legjobb megadását a szinguláris felbontásról szóló állítás harmadik verziójában szereplő $\mathbf{u}_1, \dots, \mathbf{u}_p$ és $\mathbf{v}_1, \dots, \mathbf{v}_q$ ortonormált vektorpár segítségével adhatjuk meg. Valójában csak az $(\mathbf{u}_k, \mathbf{v}_k)$, $1 \le k \le s$, párokra van szükségünk az $\alpha(\mathbf{x}, \mathbf{y})$ bilineáris függvény megadásához. Ezek hasonló szerepet játszanak, mint egy szimmetrikus mátrix nem nulla sajátértékű sajátvektorai. Továbbá ezeket az $(\mathbf{u}_k, \mathbf{v}_k)$, $1 \le k \le s$, párokat hasonló szélsőérték vagy minimax feladatok megoldásaként lehet jellemezni, mint a szimmetrikus mátrixok sajátvektorait.

Nem adom meg ezen eredmények bizonyítását, csak leírom, hogy hogyan szól az a szinguláris felbontásról szóló állítás harmadik verziójából következő képlet, amelynek a segítségével ezeket a szélsőérték jellemzéseket megtalálhatjuk és bebizonyíthatjuk.

Adva egy $\mathbf{x} \in R^p$ és $\mathbf{y} \in R^q$ vektor, írjuk ezeket $\mathbf{x} = \sum_{j=1}^p a_j \mathbf{u}_j$ és $\mathbf{y} = \sum_{j=1}^q b_j \mathbf{v}_j$ alakban. Ekkor $\alpha(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y} = \sum_{j=1}^s \lambda_j a_j b_j$.

Érdemes még a következő észrevételt tenni, amelyik egyszerűen adódik a Cauchy–Schwarz egyenlőtlenségből:

Ha $\mathbf{x} = \sum\limits_{j=1}^{p} a_j \mathbf{u}_j$, akkor $\sup\limits_{\mathbf{y} \in R^q, \|\mathbf{y}\|=1} \mathbf{x}^T \mathbf{A} \mathbf{y} = \mathbf{x}^T \mathbf{A} \mathbf{y}_0$, ahol $\mathbf{y}_0 = \frac{\sum_{j=1}^{s} a_j \lambda_j \mathbf{v}_j}{\sqrt{\sum_{j=1}^{s} a_j^2 \lambda_j^2}}$, és a szuprémum értéke $\sqrt{\sum_{j=1}^{s} a_j^2 \lambda_j^2}$.

Elég az $\mathbf{u}_k$ és $\mathbf{v}_k$, $1 \le k \le s$, vektorokat megtalálni, mert azok meghatározzák a $\lambda_k = \alpha(\mathbf{u}_k, \mathbf{v}_k)$, $1 \le k \le s$, konstansokat.

A keresett $(\mathbf{u}_k, \mathbf{v}_k)$ *párok első jellemzése:* Definiáljuk az

$$\mathbf{U}_0 = \{\mathbf{u}\colon\ \mathbf{u} \in R^p, \quad \|\mathbf{u}\| = 1\} \text{ és } \mathbf{V}_0 = \{\mathbf{v}\colon\ \mathbf{v} \in R^q, \quad \|\mathbf{v}\| = 1\},$$

valamint (az $\mathbf{u}_j$, $\mathbf{v}_j$, $1 \le j \le k$, vektorok ismeretében) az

$$\mathbf{U}_k = \{\mathbf{u}\colon\ \mathbf{u} \in R^p, \quad \|\mathbf{u}\| = 1, \quad \mathbf{u}^T \mathbf{u}_j = 0,\ 1 \le j \le k\},$$

$$\mathbf{V}_k = \{\mathbf{v}\colon\ \mathbf{v} \in R^q \quad \|\mathbf{v}\| = 1, \quad \mathbf{v}^T \mathbf{v}_j = 0,\ 1 \le j \le k\}$$

halmazokat, $1 \le k \le s$. Ekkor

$$\alpha(\mathbf{u}_k, \mathbf{v}_k) = \sup_{\mathbf{u} \in \mathbf{U}_{k-1},\ \mathbf{v} \in \mathbf{V}_{k-1}} \alpha(\mathbf{u}, \mathbf{v}),$$

illetve

$$\alpha(\mathbf{u}_k, \mathbf{v}_k) = \sup_{\mathbf{u} \in \mathbf{U}_{k-1},\ \mathbf{v} \in \mathbf{V}_0} \alpha(\mathbf{u}, \mathbf{v})$$

minden $1 \le k \le s$ indexre. Továbbá a szuprémum felvétetik a keresett $(\mathbf{u}_k, \mathbf{v}_k)$ helyen.

A keresett $(\mathbf{u}_k, \mathbf{v}_k)$ *párok második jellemzése:* Definiáljuk az

$$\mathcal{U}_k = \{\mathbf{U}\colon\ \mathbf{U} \subset R^p, \quad \mathbf{U}\ k\text{-dimenziós Euklideszi tér}\}$$

halmazokat minden $1 \le k \le p$ indexre. Ekkor

$$\alpha(\mathbf{u}_k, \mathbf{v}_k) = \inf_{\mathbf{U} \in \mathcal{U}_{p-k}} \sup_{\mathbf{u} \in \mathbf{U},\ \|\mathbf{u}\|=1,\ \mathbf{v} \in R^q,\ \|\mathbf{v}\|=1} \alpha(\mathbf{u}, \mathbf{v})$$

minden $1 \le k \le s$ indexre. Továbbá a keresett $(\mathbf{u}_k, \mathbf{v}_k)$ pár megoldása ennek a minimax feladatnak.

Komplex értékű bilineáris függvények esetében hasonló eredmény érvényes. De ott az $(\mathbf{u}_k, \mathbf{v}_k)$, $1 \le k \le s$, párok meghatározásánál az $\alpha(\mathbf{u}, \mathbf{v})$ függvény helyett az $|\alpha(\mathbf{u}, \mathbf{v})|$ függvény szélsőértékét keressük. Ez egy 1 alszolút értékű komplex szám szorzó erejéig határozza meg az $\mathbf{u}_k$ és $\mathbf{v}_k$ vektorokat. Ezek megtalálása után a $\mathbf{u}_k$ vektor értékét rögzítjük, a $\mathbf{v}_k$ vektort pedig beszorozzuk egy 1 abszolút értékű komplex számmal úgy, hogy $\alpha(\mathbf{u}_k, \mathbf{v}_k)$ pozitív valós szám legyen.