Kapitola 10

Ortogonální projekce a metoda nejmenších čtverců

V této kapitole budeme pracovat pouze nad tělesem reálných čísel.

Tvrzení 10.1 *Pro každou matici* $\mathbf{A}$ *typu* $m \times n$ *platí*

1. $r(\mathbf{A}^T\mathbf{A}) = r(\mathbf{A}^T\mathbf{A}) = r(\mathbf{A})$,
2. $\mathbf{S}(\mathbf{A}^T\mathbf{A}) = \mathbf{S}(\mathbf{A}^T)$,
3. $\mathbf{R}(\mathbf{A}^T\mathbf{A}) = \mathbf{R}(\mathbf{A})$,
4. $\mathbf{N}(\mathbf{A}^T\mathbf{A}) = \mathbf{N}(\mathbf{A})$,
5. $\mathbf{N}(\mathbf{A}^T\mathbf{A}) = \mathbf{N}(\mathbf{A}^T)$.

Tvrzení 10.2 *Buď* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *soustava* m *lineárních rovnic o* n *neznámých. Pak*

1. *soustava* $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ *je vždy řešitelná,*
2. *je-li* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *řešitelná, pak množina všech řešení soustavy* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *se rovná množině všech řešení soustavy* $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$,
3. *soustava* $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ *má jednoznačné řešení právě tehdy když* $r(\mathbf{A}) = n$, *v takovém případě pak* $\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$.

Ve Větě 10.10 si ukážeme, že v případě neřešitelné soustavy rovnic $\mathbf{A}\mathbf{x} = \mathbf{b}$ každé řešení $\mathbf{x}$ soustavy $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ minimalizuje normu $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|$. Jinak řečeno, v případě neřešitelné soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ je vektor $\mathbf{A}\mathbf{x}$ nejblíže vektoru pravých stran $\mathbf{b}$ právě pro ty vektory $\mathbf{x}$, které jsou řešením soustavy $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$.

Definice 10.1 *Je-li* $\mathbf{Ax} = \mathbf{b}$ *soustava lineárních rovnic, pak soustavu* $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ *nazýváme* asociovaný systém normálních rovnic *k soustavě* $\mathbf{Ax} = \mathbf{b}$.

Příklad 10.1 Výpočet dráhy rakety.

Příklad 10.2 Známe-li QR-rozklad $\mathbf{A} = \mathbf{QR}$ matice $\mathbf{A}$ s lineárně nezávislými sloupci, můžeme jej použít k řešení soustavy $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$. Připomeňme si, že sloupce $\mathbf{Q}$ tvoří ortonormální množinu a $\mathbf{R}$ je horní trojúhelníková matice s kladnými prvky na hlavní diagonále. Potom platí

$$\mathbf{A}^T\mathbf{A} = \mathbf{R}^T\mathbf{Q}^T\mathbf{QR} = \mathbf{R}^T\mathbf{R}.$$

Protože $\mathbf{Q}$ má ortonormální sloupce, platí $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ a soustava $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ je tak ekvivalentní soustavě $\mathbf{R}^T\mathbf{Rx} = \mathbf{R}^T\mathbf{Q}^T\mathbf{b}$. Matice $\mathbf{R}^T$ je regulární, takže tato soustava je dále ekvivalentní soustavě $\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$, kterou můžeme vyřešit zpětnou substitucí a dostaneme tak jediné řešení $\mathbf{x} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}$ soustavy $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$.

Věta 10.3 Věta o ortogonálním rozkladu. *Pro každou matici* $\mathbf{A}$ *typu* $m \times n$ *nad* $\mathbf{R}$ *platí*

1. $\mathbf{S}(\mathbf{A})^\perp = \mathbf{N}(\mathbf{A}^T)$,
2. $\mathbf{N}(\mathbf{A})^\perp = \mathbf{S}(\mathbf{A}^T)$.

Důsledek 10.4 Fredholmova alternativa. *Pro každou matici* $\mathbf{A}$ *nastává právě jedna z následujících dvou možností:*

1. *soustava* $\mathbf{Ax} = \mathbf{b}$ *má řešení pro každou pravou stranu* $\mathbf{b}$,
2. *soustava* $\mathbf{A}^T\mathbf{y} = \mathbf{0}$ *má nenulové řešení.*

Tvrzení 10.5 *Buď* $\mathbf{A}$ *matice typu* $m \times n$ *nad* $\mathbf{R}$ *a* $r(\mathbf{A}) = r$. *Pak existují ortogonální matice* $\mathbf{U}$ *řádu* m, *ortogonální matice* $\mathbf{V}$ *řádu* n *a regulární matice* $\mathbf{C}$ *řádu* r *takové, že*

$$\mathbf{A} = \mathbf{U}\begin{pmatrix} \mathbf{C} & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}\mathbf{V}^T.$$

Prvních r *sloupců matice* $\mathbf{U}$ *tvoří ortonormální bázi* $\mathbf{S}(\mathbf{A})$, *posledních* $m-r$ *sloupců tvoří ortonormální bázi* $\mathbf{N}(\mathbf{A}^T)$. *Prvních* r *sloupců* $\mathbf{V}$ *tvoří ortonormální bázi* $\mathbf{S}(\mathbf{A}^T)$ *a posledních* $n-r$ *sloupců tvoří ortonormální bázi* $\mathbf{N}(\mathbf{A})$.

Definice 10.2 *Buď* $\mathbf{A}$ *matice typu* $m \times n$ *nad* $\mathbf{R}$ *a* $r(\mathbf{A}) = r$. *Vyjádření* $\mathbf{A} = \mathbf{URV}^T$, *kde* $\mathbf{U}$ *je ortogonální matice řádu* m, $\mathbf{V}$ *je ortogonální matice řádu* n *a*

$$\mathbf{R} = \begin{pmatrix} \mathbf{C} & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}$$

je matice typu $m\times n$, *a* $\mathbf{C}$ *je regulární matice řádu* r, *se nazývá* URV-rozklad matice $\mathbf{A}$.

Všimněte si, že URV-rozklad matice $\mathbf{A}$ není určený jednoznačně.

Tvrzení 10.6 *Buď* $\mathbf{A} = \mathbf{URV}^T$ *URV-rozklad matice* $\mathbf{A}$. *Matice*

$$\mathbf{B} = \mathbf{V} \begin{pmatrix} \mathbf{C}^{-1} & \mathbf{0}_{r\times(m-r)} \\ \mathbf{0}_{(n-r)\times r} & \mathbf{0}_{(n-r)\times(m-r)} \end{pmatrix} \mathbf{U}^T,$$

splňuje následující rovnosti

1. $\mathbf{ABA} = \mathbf{A}$,
2. $\mathbf{BAB} = \mathbf{B}$,
3. $(\mathbf{AB})^T = \mathbf{AB}$,
4. $(\mathbf{BA})^T = \mathbf{BA}$.

Matice $\mathbf{B}$ *je těmito čtyřmi rovnostmi určena jednoznačně.*

Definice 10.3 *Buď* $\mathbf{A}$ *matice typu* $m \times n$ *nad* $\mathbf{R}$. *Matici* $\mathbf{B}$ *splňující rovnosti 1,2,3,4 z Tvrzení 10.6 nazýváme* Moore-Penroseova inverze *nebo také* pseudoinverze *matice* $\mathbf{A}$ *a označujeme ji* $\mathbf{A}^\dagger$.

Tvrzení 10.7 *Buď* $\mathbf{M}$ *podprostor* $\mathbf{R}^n$ *a* $\mathbf{M}^\perp$ *jeho ortogonální doplněk. Potom každý vektor* $\mathbf{x} \in \mathbf{R}^n$ *lze vyjádřit právě jedním způsobem ve tvaru* $\mathbf{x} = \mathbf{p} + \mathbf{q}$, *kde* $\mathbf{p} \in \mathbf{M}$ *a* $\mathbf{q} \in \mathbf{M}^\perp$.

Definice 10.4 *Je-li* $\mathbf{M}$ *podprostor* $\mathbf{R}^n$ *a* $\mathbf{x} = \mathbf{p} + \mathbf{q}$, *kde* $\mathbf{p} \in \mathbf{M}$ *a* $\mathbf{q} \in \mathbf{M}^\perp$, *pak vektor* $\mathbf{p}$ *nazýváme* ortogonální projekce $\mathbf{x}$ *na podprostor* $\mathbf{M}$.

Tvrzení 10.8 *Buď* $\mathbf{M}$ *podprostor* $\mathbf{R}^n$, $\mathbf{x} \in \mathbf{R}^n$ *a* $\mathbf{p} \in \mathbf{M}$ *ortogonální projekce vektoru* $\mathbf{x}$ *na* $\mathbf{M}$. *Pak pro každý vektor* $\mathbf{m} \in \mathbf{M}$ *platí*

$$\|\mathbf{x} - \mathbf{p}\| \le \|\mathbf{x} - \mathbf{m}\|,$$

přičemž rovnost nastává právě když $\mathbf{m} = \mathbf{p}$.

Tvrzení 10.9 *Buď* $\mathbf{A}$ *matice typu* $m \times n$ *a* $\mathbf{A}^\dagger$ *Moore-Penroseova inverze* $\mathbf{A}$. *Potom se ortogonální projekce libovolného vektoru* $\mathbf{b} \in \mathbf{R}^m$ *na sloupcový prostor* $\mathbf{S}(\mathbf{A})$ *matice* $\mathbf{A}$ *rovná* $(\mathbf{A}\mathbf{A}^\dagger)\mathbf{b}$.

Věta 10.10 Metoda nejmenších čtverců *Nechť* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *je soustava lineárních rovnic nad* $\mathbf{R}$. *Pak pro vektor* $\mathbf{x} \in \mathbf{R}^n$ *je ekvivalentní*

1. $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$,
2. $\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{A}^\dagger)\mathbf{b}$.