

Γραμμική Άλγεβρα

Εικοστή Διαλεξη

Βασίλειος Νάκος

Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστήμιο Αθηνών

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle}$$

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} =$$

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} = \sqrt{u^T u}$$

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} = \sqrt{u^T u} \text{ ως το γεωμετρικό του μήκος.}$$

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} = \sqrt{u^T u} \text{ ως το γεωμετρικό του μήκος.}$$

Έστω πίνακας $U \in \mathbb{R}^{n \times n}$ ώστε $U^T U = I$. Τότε,

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} = \sqrt{u^T u} \text{ ως το γεωμετρικό του μήκος.}$$

Έστω πίνακας $U \in \mathbb{R}^{n \times n}$ ώστε $U^T U = I$. Τότε,

- ο U^T ανάστροφος είναι ο αντίστροφος του U (άρα και $UU^T = I$)

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} = \sqrt{u^T u} \text{ ως το γεωμετρικό του μήκος.}$$

Έστω πίνακας $U \in \mathbb{R}^{n \times n}$ ώστε $U^T U = I$. Τότε,

- ο U^T ανάστροφος είναι ο αντίστροφος του U (άρα και $U U^T = I$),
- οι στήλες και οι γραμμές του U είναι ορθοκανονικές (κάθε στήλη γεωμετρικό μήκος 1 και όλες κάθετες μεταξύ τους),

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} = \sqrt{u^T u} \text{ ως το γεωμετρικό του μήκος.}$$

Έστω πίνακας $U \in \mathbb{R}^{n \times n}$ ώστε $U^T U = I$. Τότε,

- ο U^T ανάστροφος είναι ο αντίστροφος του U (άρα και $U U^T = I$),
- οι στήλες και οι γραμμές του U είναι ορθοκανονικές (κάθε στήλη γεωμετρικό μήκος 1 και όλες κάθετες μεταξύ τους),
- και πιο σημαντικά, $\|Ux\|_2^2 = \|x\|_2^2$, δηλαδή ο U αφήνει ίδιο το γεωμετρικό μήκος του x .

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} = \sqrt{u^T u} \text{ ως το γεωμετρικό του μήκος.}$$

Έστω πίνακας $U \in \mathbb{R}^{n \times n}$ ώστε $U^T U = I$. Τότε,

- ο U^T ανάστροφος είναι ο αντίστροφος του U (άρα και $U U^T = I$),
- οι στήλες και οι γραμμές του U είναι ορθοκανονικές (κάθε στήλη γεωμετρικό μήκος 1 και όλες κάθετες μεταξύ τους),
- και πιο σημαντικά, $\|Ux\|_2^2 = \|x\|_2^2$, δηλαδή ο U αφήνει ίδιο το γεωμετρικό μήκος του x .

Λέμε ότι ο U είναι **ορθοκανονικός πίνακας**

Πίνακες ως στροφές.

Έστω διάνυσμα $u \in \mathbb{R}^n$. Ορίζουμε

$$\|u\|_2 := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n u_i^2} = \sqrt{u^T u} \text{ ως το γεωμετρικό του μήκος.}$$

Έστω πίνακας $U \in \mathbb{R}^{n \times n}$ ώστε $U^T U = I$. Τότε,

- ο U^T ανάστροφος είναι ο αντίστροφος του U (άρα και $U U^T = I$),
- οι στήλες και οι γραμμές του U είναι ορθοκανονικές (κάθε στήλη γεωμετρικό μήκος 1 και όλες κάθετες μεταξύ τους),
- και πιο σημαντικά, $\|Ux\|_2^2 = \|x\|_2^2$, δηλαδή ο U αφήνει ίδιο το γεωμετρικό μήκος του x .

Λέμε ότι ο U είναι **ορθοκανονικός πίνακας** και ότι η απεικόνιση $x \rightarrow Ux$ στρέφει το διάνυσμα x .

Οι παρακάτω πίνακες είναι ορθοκανονικοί και αντιστοιχούν σε ταυτοτική απεικόνιση, σε στροφή κατά γωνία θ , σε ανάκλαση κατά τον άξονα των x και αναδιάταξη των συντεταγμένων.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Οι παρακάτω πίνακες είναι ορθοκανονικοί και αντιστοιχούν σε ταυτοτική απεικόνιση, σε στροφή κατά γωνία θ , σε ανάκλαση κατά τον άξονα των x και αναδιάταξη των συντεταγμένων.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Αν U ορθοκανονικός τότε η απεικόνιση $x \rightarrow Ux$ *στρίβει* το x :

Οι παρακάτω πίνακες είναι ορθοκανονικοί και αντιστοιχούν σε ταυτοτική απεικόνιση, σε στροφή κατά γωνία θ , σε ανάκλαση κατά τον άξονα των x και αναδιάταξη των συντεταγμένων.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Αν U ορθοκανονικός τότε η απεικόνιση $x \rightarrow Ux$ *στρίβει* το x :

$$U^T U = I \Rightarrow \|Ux\|_2^2 =$$

Οι παρακάτω πίνακες είναι ορθοκανονικοί και αντιστοιχούν σε ταυτοτική απεικόνιση, σε στροφή κατά γωνία θ , σε ανάκλαση κατά τον άξονα των x και αναδιάταξη των συντεταγμένων.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Αν U ορθοκανονικός τότε η απεικόνιση $x \rightarrow Ux$ *στρίβει* το x :

$$U^T U = I \Rightarrow \|Ux\|_2^2 = (Ux)^T (Ux) =$$

Οι παρακάτω πίνακες είναι ορθοκανονικοί και αντιστοιχούν σε ταυτοτική απεικόνιση, σε στροφή κατά γωνία θ , σε ανάκλαση κατά τον άξονα των x και αναδιάταξη των συντεταγμένων.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Αν U ορθοκανονικός τότε η απεικόνιση $x \rightarrow Ux$ *στρίβει* το x :

$$U^T U = I \Rightarrow \|Ux\|_2^2 = (Ux)^T (Ux) = x^T \underbrace{U^T U}_I x =$$

Οι παρακάτω πίνακες είναι ορθοκανονικοί και αντιστοιχούν σε ταυτοτική απεικόνιση, σε στροφή κατά γωνία θ , σε ανάκλαση κατά τον άξονα των x και αναδιάταξη των συντεταγμένων.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Αν U ορθοκανονικός τότε η απεικόνιση $x \rightarrow Ux$ *στρίβει* το x :

$$U^T U = I \Rightarrow \|Ux\|_2^2 = (Ux)^T (Ux) = x^T \underbrace{U^T U}_I x = x^T x =$$

Οι παρακάτω πίνακες είναι ορθοκανονικοί και αντιστοιχούν σε ταυτοτική απεικόνιση, σε στροφή κατά γωνία θ , σε ανάκλαση κατά τον άξονα των x και αναδιάταξη των συντεταγμένων.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Αν U ορθοκανονικός τότε η απεικόνιση $x \rightarrow Ux$ *στρίβει* το x :

$$U^T U = I \Rightarrow \|Ux\|_2^2 = (Ux)^T (Ux) = x^T \underbrace{U^T U}_I x = x^T x = \|x\|_2^2.$$

Αποσύνθεση κάθε γραμμικής απεικόνισης σε τρεις πράξεις.

Κάθε γραμμική απεικόνιση $x \rightarrow Ax$ μπορούμε να την δούμε ως σύνθεση τριών απεικονίσεων.

Θεώρημα

Έστω $A \in \mathbb{R}^{m \times n}$. Τότε υπάρχει ένας ορθοκανονικός πίνακας $U \in \mathbb{R}^{m \times m}$, ένας ορθοκανονικός πίνακας $V \in \mathbb{R}^{n \times n}$ και ένας πίνακας $\Sigma \in \mathbb{R}^{m \times n}$ ο οποίος έχει μη μηδενικά στοιχεία πάνω στη 'διαγώνιο' ώστε

$$A = U\Sigma V^T.$$

Αποσύνθεση κάθε γραμμικής απεικόνισης σε τρεις πράξεις.

Κάθε γραμμική απεικόνιση $x \rightarrow Ax$ μπορούμε να την δούμε ως σύνθεση τριών απεικονίσεων.

Θεώρημα

Έστω $A \in \mathbb{R}^{m \times n}$. Τότε υπάρχει ένας ορθοκανονικός πίνακας $U \in \mathbb{R}^{m \times m}$, ένας ορθοκανονικός πίνακας $V \in \mathbb{R}^{n \times n}$ και ένας πίνακας $\Sigma \in \mathbb{R}^{m \times n}$ ο οποίος έχει μη μηδενικά στοιχεία πάνω στη 'διαγώνιο' ώστε

$$A = U\Sigma V^T.$$

Λέμε την αποσύνθεση $A = U\Sigma V^T$ αποσύνθεση ιδιάζουσων τιμών.

Για παράδειγμα για $A \in \mathbb{R}^{4 \times 3}$ έχουμε

$$A = \underbrace{\begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{bmatrix}}_U \cdot \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{bmatrix}}_\Sigma \underbrace{\begin{bmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{bmatrix}}_{V^T}$$

Άρα τι συμβαίνει στο x όταν εφαρμόζουμε τον A ;

Για παράδειγμα για $A \in \mathbb{R}^{4 \times 3}$ έχουμε

$$A = \underbrace{\begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{bmatrix}}_U \cdot \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{bmatrix}}_\Sigma \underbrace{\begin{bmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{bmatrix}}_{V^T}$$

Άρα τι συμβαίνει στο x όταν εφαρμόζουμε τον A ;

- 1 Η απεικόνιση $x \rightarrow V^T x$ στρίβει το x και μας δίνει το $y = V^T x$,

Για παράδειγμα για $A \in \mathbb{R}^{4 \times 3}$ έχουμε

$$A = \underbrace{\begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{bmatrix}}_U \cdot \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{bmatrix}}_\Sigma \underbrace{\begin{bmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{bmatrix}}_{V^T}$$

Άρα τι συμβαίνει στο x όταν εφαρμόζουμε τον A ;

- 1 Η απεικόνιση $x \rightarrow V^T x$ στρίβει το x και μας δίνει το $y = V^T x$,
- 2 η απεικόνιση $y \rightarrow \Sigma y$ διαστέλλει κάθε συντεταγμένη του y (την i -οστή κατά σ_i) και βάζει άλλες $m - n$ μηδενικές συντεταγμένες δίνοντας μας το $z = \Sigma y$,

Για παράδειγμα για $A \in \mathbb{R}^{4 \times 3}$ έχουμε

$$A = \underbrace{\begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{bmatrix}}_U \cdot \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{bmatrix}}_\Sigma \underbrace{\begin{bmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{bmatrix}}_{V^T}$$

Άρα τι συμβαίνει στο x όταν εφαρμόζουμε τον A ;

- 1 Η απεικόνιση $x \rightarrow V^T x$ στρίβει το x και μας δίνει το $y = V^T x$,
- 2 η απεικόνιση $y \rightarrow \Sigma y$ διαστέλλει κάθε συντεταγμένη του y (την i -οστή κατά σ_i) και βάζει άλλες $m - n$ μηδενικές συντεταγμένες δίνοντας μας το $z = \Sigma y$,
- 3 η απεικόνιση $z \rightarrow Uz$ στρίβει το z (στον $\mathbb{R}^m$) και άρα παίρνουμε το $Ax = U\Sigma V^T x$.

Άλλες ιδιότητες της αποσύνθεσης.

- Οι διαγώνιες τιμές $\sigma_1, \sigma_2, \dots$ ικανοποιούν $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_n \geq 0$.

Άλλες ιδιότητες της αποσύνθεσης.

- Οι διαγώνιες τιμές $\sigma_1, \sigma_2, \dots$ ικανοποιούν $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_n \geq 0$. Οι τιμές αυτές λέγονται *ιδιάζουσες τιμές*.
- Αν ο βαθμός του A είναι r τότε έχουμε ακριβώς r μη μηδενικές ιδιάζουσες τιμές,

Άλλες ιδιότητες της αποσύνθεσης.

- Οι διαγώνιες τιμές $\sigma_1, \sigma_2, \dots$ ικανοποιούν $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_n \geq 0$. Οι τιμές αυτές λέγονται *ιδιάζουσες τιμές*.
- Αν ο βαθμός του A είναι r τότε έχουμε ακριβώς r μη μηδενικές ιδιάζουσες τιμές,
- Οι πρώτες r στήλες του U είναι μια βάση του χώρου στηλών του A ,

Άλλες ιδιότητες της αποσύνθεσης.

- Οι διαγώνιες τιμές $\sigma_1, \sigma_2, \dots$ ικανοποιούν $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_n \geq 0$. Οι τιμές αυτές λέγονται *ιδιάζουσες τιμές*.
- Αν ο βαθμός του A είναι r τότε έχουμε ακριβώς r μη μηδενικές ιδιάζουσες τιμές,
- Οι πρώτες r στήλες του U είναι μια βάση του χώρου στηλών του A ,
- Οι τελευταίες $n - r$ στήλες του V είναι μια βάση του μηδενοχώρου του A .

Άλλες ιδιότητες της αποσύνθεσης.

- Οι διαγώνιες τιμές $\sigma_1, \sigma_2, \dots$ ικανοποιούν $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_n \geq 0$. Οι τιμές αυτές λέγονται *ιδιάζουσες τιμές*.
- Αν ο βαθμός του A είναι r τότε έχουμε ακριβώς r μη μηδενικές ιδιάζουσες τιμές,
- Οι πρώτες r στήλες του U είναι μια βάση του χώρου στηλών του A ,
- Οι τελευταίες $n - r$ στήλες του V είναι μια βάση του μηδενοχώρου του A .

Πως βρίσκουμε τις ιδιάζουσες τιμές του A ;

Τα σ_i^2 είναι οι ιδιοτιμές του συμμετρικού πίνακα $A^T A$

Άλλες ιδιότητες της αποσύνθεσης.

- Οι διαγώνιες τιμές $\sigma_1, \sigma_2, \dots$ ικανοποιούν $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_n \geq 0$. Οι τιμές αυτές λέγονται *ιδιάζουσες τιμές*.
- Αν ο βαθμός του A είναι r τότε έχουμε ακριβώς r μη μηδενικές ιδιάζουσες τιμές,
- Οι πρώτες r στήλες του U είναι μια βάση του χώρου στηλών του A ,
- Οι τελευταίες $n - r$ στήλες του V είναι μια βάση του μηδενοχώρου του A .

Πως βρίσκουμε τις ιδιάζουσες τιμές του A ;

Τα σ_i^2 είναι οι ιδιοτιμές του συμμετρικού πίνακα $A^T A = V(\Sigma^T \Sigma)V^T$,

Άλλες ιδιότητες της αποσύνθεσης.

- Οι διαγώνιες τιμές $\sigma_1, \sigma_2, \dots$ ικανοποιούν $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_n \geq 0$. Οι τιμές αυτές λέγονται *ιδιάζουσες τιμές*.
- Αν ο βαθμός του A είναι r τότε έχουμε ακριβώς r μη μηδενικές ιδιάζουσες τιμές,
- Οι πρώτες r στήλες του U είναι μια βάση του χώρου στηλών του A ,
- Οι τελευταίες $n - r$ στήλες του V είναι μια βάση του μηδενοχώρου του A .

Πως βρίσκουμε τις ιδιάζουσες τιμές του A ;

Τα σ_i^2 είναι οι ιδιοτιμές του συμμετρικού πίνακα $A^T A = V(\Sigma^T \Sigma)V^T$, καθώς η γραφή $\underbrace{V}_P \underbrace{(\Sigma^T \Sigma)}_\Lambda \underbrace{V^T}_{P^{-1}}$ είναι μια διαγωνιοποίηση του $A^T A$.

Προβολή πάνω σε χώρο/ γραμμική παλινδρόμηση.

Έστω $A \in \mathbb{R}^{m \times n}$ και $b \in \mathbb{R}^m$.

$$x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2^2.$$

Προβολή πάνω σε χώρο/ γραμμική παλινδρόμηση.

Έστω $A \in \mathbb{R}^{m \times n}$ και $b \in \mathbb{R}^m$.

$$x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2^2.$$

$\min_{x_1, x_2} (x_1 + x_2 - 1)^2 + (x_1 - x_2)^2 + (x_1 + 9x_2 - 100)^2$ ισοδύναμο με

$$\left\| \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 9 \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_x - \underbrace{\begin{bmatrix} 1 \\ 0 \\ 100 \end{bmatrix}}_b \right\|_2^2$$

- Εφαρμογή: Έχουμε ένα ενδεχομένως αδύνατο σύστημα $Ax = b$, αλλά θέλω να βρούμε μία λύση x η οποία είναι κοντά στο b .
- Το πρόβλημα $x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ μας ζητάει να προβάλουμε το b πάνω στον χώρο στηλών του A
 $\{Ax \mid x \in \mathbb{R}^n\} = \operatorname{span}\{A_1, \dots, A_n\}$.

- Εφαρμογή: Έχουμε ένα ενδεχομένως αδύνατο σύστημα $Ax = b$, αλλά θέλω να βρούμε μία λύση x η οποία είναι κοντά στο b .
- Το πρόβλημα $x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ μας ζητάει να προβάλουμε το b πάνω στον χώρο στηλών του A
 $\{Ax \mid x \in \mathbb{R}^n\} = \operatorname{span}\{A_1, \dots, A_n\}$.
- Η τιμή x^* αντιστοιχεί στον γραμμικός συνδυασμός εκείνος των στηλών του A ο οποίος πιάνει το ίχνος του b στον χώρο στηλών του A .

- Εφαρμογή: Έχουμε ένα ενδεχομένως αδύνατο σύστημα $Ax = b$, αλλά θέλω να βρούμε μία λύση x η οποία είναι κοντά στο b .
- Το πρόβλημα $x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ μας ζητάει να προβάλουμε το b πάνω στον χώρο στηλών του A
 $\{Ax \mid x \in \mathbb{R}^n\} = \operatorname{span}\{A_1, \dots, A_n\}$.
- Η τιμή x^* αντιστοιχεί στον γραμμικός συνδυασμός εκείνος των στηλών του A ο οποίος πιάνει το ίχνος του b στον χώρο στηλών του A .
- $b = b^\perp + b^\parallel$

- Εφαρμογή: Έχουμε ένα ενδεχομένως αδύνατο σύστημα $Ax = b$, αλλά θέλω να βρούμε μία λύση x η οποία είναι κοντά στο b .
- Το πρόβλημα $x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ μας ζητάει να προβάλουμε το b πάνω στον χώρο στηλών του A
 $\{Ax \mid x \in \mathbb{R}^n\} = \operatorname{span}\{A_1, \dots, A_n\}$.
- Η τιμή x^* αντιστοιχεί στον γραμμικός συνδυασμός εκείνος των στηλών του A ο οποίος πιάνει το ίχνος του b στον χώρο στηλών του A .
- $b = b^\perp + b^\parallel$ και $b^\parallel = Ax^*$ και

- Εφαρμογή: Έχουμε ένα ενδεχομένως αδύνατο σύστημα $Ax = b$, αλλά θέλω να βρούμε μία λύση x η οποία είναι κοντά στο b .
- Το πρόβλημα $x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ μας ζητάει να προβάλουμε το b πάνω στον χώρο στηλών του A
 $\{Ax \mid x \in \mathbb{R}^n\} = \operatorname{span}\{A_1, \dots, A_n\}$.
- Η τιμή x^* αντιστοιχεί στον γραμμικός συνδυασμός εκείνος των στηλών του A ο οποίος πιάνει το ίχνος του b στον χώρο στηλών του A .
- $b = b^\perp + b^\parallel$ και $b^\parallel = Ax^*$ και $\|Ax^* - b\|_2^2 = \|b^\perp\|_2^2$.

- Εφαρμογή: Έχουμε ένα ενδεχομένως αδύνατο σύστημα $Ax = b$, αλλά θέλω να βρούμε μία λύση x η οποία είναι κοντά στο b .
- Το πρόβλημα $x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ μας ζητάει να προβάλουμε το b πάνω στον χώρο στηλών του A
 $\{Ax \mid x \in \mathbb{R}^n\} = \operatorname{span}\{A_1, \dots, A_n\}$.
- Η τιμή x^* αντιστοιχεί στον γραμμικός συνδυασμός εκείνος των στηλών του A ο οποίος πιάνει το ίχνος του b στον χώρο στηλών του A .
- $b = b^\perp + b^\parallel$ και $b^\parallel = Ax^*$ και $\|Ax^* - b\|_2^2 = \|b^\perp\|_2^2$.

Μπορούμε να χρησιμοποιήσουμε την αποσύνθεση σε ιδιάζουσες τιμές του A για να βρούμε το $b^\parallel$ (ίχνος πάνω στον χώρο στηλών), το x^* και το $b^\perp$.

$$\|Ax - b\|_2^2 =$$

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 =$$

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

εφόσον $\|z\|_2^2 = \|U^T z\|_2^2, \forall z \in \mathbb{R}^m$.

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

εφόσον $\|z\|_2^2 = \|U^T z\|_2^2, \forall z \in \mathbb{R}^m$. Επιπλέον,

$$\Sigma = \begin{bmatrix} \tilde{\Sigma} \\ 0_{(m-n) \times n} \end{bmatrix}$$

όπου $\tilde{\Sigma}$ ο πίνακας που έχει στη διαγώνιο τις ιδιάζουσες τιμές,

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

εφόσον $\|z\|_2^2 = \|U^T z\|_2^2, \forall z \in \mathbb{R}^m$. Επιπλέον,

$$\Sigma = \begin{bmatrix} \tilde{\Sigma} \\ 0_{(m-n) \times n} \end{bmatrix}$$

όπου $\tilde{\Sigma}$ ο πίνακας που έχει στη διαγώνιο τις ιδιάζουσες τιμές, και άρα έχουμε

$$\|\Sigma V^T x - U^T b\|_2^2 =$$

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

εφόσον $\|z\|_2^2 = \|U^T z\|_2^2, \forall z \in \mathbb{R}^m$. Επιπλέον,

$$\Sigma = \begin{bmatrix} \tilde{\Sigma} \\ 0_{(m-n) \times n} \end{bmatrix}$$

όπου $\tilde{\Sigma}$ ο πίνακας που έχει στη διαγώνιο τις ιδιάζουσες τιμές, και άρα έχουμε

$$\|\Sigma V^T x - U^T b\|_2^2 = \left\| \begin{bmatrix} \tilde{\Sigma}(V^T x) - (U^T b)_{[1, \dots, n]} \\ -(U^T b)_{[n+1, \dots, m]} \end{bmatrix} \right\|_2^2$$

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

εφόσον $\|z\|_2^2 = \|U^T z\|_2^2, \forall z \in \mathbb{R}^m$. Επιπλέον,

$$\Sigma = \begin{bmatrix} \tilde{\Sigma} \\ 0_{(m-n) \times n} \end{bmatrix}$$

όπου $\tilde{\Sigma}$ ο πίνακας που έχει στη διαγώνιο τις ιδιάζουσες τιμές, και άρα έχουμε

$$\|\Sigma V^T x - U^T b\|_2^2 = \left\| \begin{bmatrix} \tilde{\Sigma}(V^T x) - (U^T b)_{[1, \dots, n]} \\ -(U^T b)_{[n+1, \dots, m]} \end{bmatrix} \right\|_2^2$$

άρα για να ελαχιστοποιηθεί το $\|Ax - b\|_2^2$ αρκεί να μηδενίσω τον $\tilde{\Sigma}(V^T x) - (U^T b)_{[1, \dots, n]}$,

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

εφόσον $\|z\|_2^2 = \|U^T z\|_2^2, \forall z \in \mathbb{R}^m$. Επιπλέον,

$$\Sigma = \begin{bmatrix} \tilde{\Sigma} \\ 0_{(m-n) \times n} \end{bmatrix}$$

όπου $\tilde{\Sigma}$ ο πίνακας που έχει στη διαγώνιο τις ιδιάζουσες τιμές, και άρα έχουμε

$$\|\Sigma V^T x - U^T b\|_2^2 = \left\| \begin{bmatrix} \tilde{\Sigma}(V^T x) - (U^T b)_{[1, \dots, n]} \\ -(U^T b)_{[n+1, \dots, m]} \end{bmatrix} \right\|_2^2$$

άρα για να ελαχιστοποιηθεί το $\|Ax - b\|_2^2$ αρκεί να μηδενίσω τον $\tilde{\Sigma}(V^T x) - (U^T b)_{[1, \dots, n]}$, και τότε πληρώνω ακριβώς $\|(U^T b)_{[n+1, \dots, m]}\|_2^2$.

$$\|Ax - b\|_2^2 = \|U(\Sigma V^T x) - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

εφόσον $\|z\|_2^2 = \|U^T z\|_2^2, \forall z \in \mathbb{R}^m$. Επιπλέον,

$$\Sigma = \begin{bmatrix} \tilde{\Sigma} \\ 0_{(m-n) \times n} \end{bmatrix}$$

όπου $\tilde{\Sigma}$ ο πίνακας που έχει στη διαγώνιο τις ιδιάζουσες τιμές, και άρα έχουμε

$$\|\Sigma V^T x - U^T b\|_2^2 = \left\| \begin{bmatrix} \tilde{\Sigma}(V^T x) - (U^T b)_{[1, \dots, n]} \\ -(U^T b)_{[n+1, \dots, m]} \end{bmatrix} \right\|_2^2$$

άρα για να ελαχιστοποιηθεί το $\|Ax - b\|_2^2$ αρκεί να μηδενίσω τον $\tilde{\Sigma}(V^T x) - (U^T b)_{[1, \dots, n]}$, και τότε πληρώνω ακριβώς $\|(U^T b)_{[n+1, \dots, m]}\|_2^2$.

Οπότε $\tilde{\Sigma}(V^T x) = (U^T b)_{[1, \dots, n]}$ και άρα $x^* = V\tilde{\Sigma}^{-1}(U^T b)_{[1, \dots, n]}$.

$$x^{\star} = \operatorname{argmin} \|Ax - b\|_2^2, \quad A = U\Sigma V^T$$

$$x^{\star} = \operatorname{argmin} \|Ax - b\|_2^2, \quad A = U\Sigma V^T$$

① $x^{\star} = V\tilde{\Sigma}^{-1}(U^T b)_{[1,\dots,n]},$

$$x^* = \operatorname{argmin} \|Ax - b\|_2^2, \quad A = U\Sigma V^T$$

- 1 $x^* = V\tilde{\Sigma}^{-1}(U^T b)_{[1,\dots,n]},$
- 2 $b^\parallel$, η προβολή του b πάνω στον χώρο στηλών του A

$$x^{\star} = \operatorname{argmin} \|Ax - b\|_2^2, \quad A = U\Sigma V^T$$

- ❶ $x^{\star} = V\tilde{\Sigma}^{-1}(U^T b)_{[1,\dots,n]},$
- ❷ $b^{\parallel}$, η προβολή του b πάνω στον χώρο στηλών του A ισούται με
 $b^{\parallel} = Ax^{\star} = U\Sigma\tilde{\Sigma}^{-1}(U^T b)_{[1,\dots,n]},$
- ❸ $b^{\perp} = b - b^{\parallel}$

$$x^* = \operatorname{argmin} \|Ax - b\|_2^2, \quad A = U\Sigma V^T$$

- ❶ $x^* = V\tilde{\Sigma}^{-1}(U^T b)_{[1,\dots,n]},$
- ❷ $b^{\parallel}$, η προβολή του b πάνω στον χώρο στηλών του A ισούται με $b^{\parallel} = Ax^* = U\Sigma\tilde{\Sigma}^{-1}(U^T b)_{[1,\dots,n]},$
- ❸ $b^{\perp} = b - b^{\parallel}$ και η απόσταση του b από τον χώρο στηλών ισούται με $\|b^{\perp}\|_2^2 = \|(U^T b)_{[n+1,\dots,m]}\|_2^2$

Τα $b^{\parallel}, b^{\perp}$ δεν εξαρτώνται από τον V^T !

$$x^* = \operatorname{argmin} \|Ax - b\|_2^2, \quad A = U\Sigma V^T$$

- ❶ $x^* = V\tilde{\Sigma}^{-1}(U^T b)_{[1,\dots,n]},$
- ❷ $b^{\parallel}$, η προβολή του b πάνω στον χώρο στηλών του A ισούται με $b^{\parallel} = Ax^* = U\Sigma\tilde{\Sigma}^{-1}(U^T b)_{[1,\dots,n]},$
- ❸ $b^{\perp} = b - b^{\parallel}$ και η απόσταση του b από τον χώρο στηλών ισούται με $\|b^{\perp}\|_2^2 = \|(U^T b)_{[n+1,\dots,m]}\|_2^2$

Τα $b^{\parallel}, b^{\perp}$ δεν εξαρτώνται από τον V^T ! Ο V^T κάνει μόνο αλλαγή βάσης επί της ουσίας καθώς ισχύει ότι ο χώρος στηλών του A και του $U\Sigma$ είναι ίδιος.

Η Γραμμική Άλγεβρα είναι (πλέον) το κυρίαρχο μαθηματικό εργαλείο της Επιστήμης Υπολογιστών.

Η Συνέχεια στο επόμενο επεισόδιο!