AT: session 1

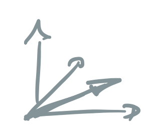
Montag, 23. April 2018 17:10

Linear Algebra Recap

1. vector $(\vec{x}, \bar{x}, \mathbf{x})$

geometrical object in some space | with magnitude | and direction

e.g. 3D | length | direction



vector space $\mathbb{V}$: collection of vector objects
$(+), (\cdot), \alpha, 1, -\vec{x}$ (w.r.t. "+")
cumulativ and associative

2. linear independence (l.i.)

if $\alpha\vec{a} + \beta\vec{b} + \dots + \delta\vec{s} = \vec{0} \Rightarrow \alpha, \beta, \dots, \delta = 0 \to$ only way to fulfill equation
then $\vec{a}, \vec{b}, \dots, \vec{s}$ are l.i

3. basis vectors $\{\vec{e}_i\} = \{\vec{e}_1 .. \vec{e}_N\}$

Say $\mathbb{V}$ is N-dimensional
$\Rightarrow$ set of N l.i. vectors $\vec{e}_i$ form a basis
$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_N\vec{e}_N = \sum_{i=1}^{N} x_i\vec{e}_i \quad \forall \vec{x} \in \mathbb{V}$

4. inner product $\vec{a}\cdot\vec{b} = (\vec{a}, \vec{b}) = \langle \vec{a} | \vec{b} \rangle$

- orthogonality: $\vec{a}\cdot\vec{b} = 0$
- norm: $\sqrt{\vec{a}\cdot\vec{b}} = \|\vec{a}\|_2$ ($\hat{=}$ length of the vector)
- orthonormal basis: $\hat{e}_i \cdot \hat{e}_j = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & \text{else} \end{cases}$

and $\|\hat{e}_i\|_2 = 1 \quad \forall i$

$\Rightarrow \vec{a} = \sum_{i=1}^{N} a_i \hat{e}_i \qquad \vec{b} = \sum_{i=1}^{N} b_i \hat{e}_i$

$\vec{a}\cdot\vec{b} = \sum_{i=1}^{N} a_i b_i = \vec{a}^T\vec{b}$

[note $\underbrace{\hat{e}_i \cdot \hat{e}_j = G_{ij}}$ $\Rightarrow \vec{a}\cdot\vec{b} = \sum_{i=1}^{N}\sum_{j=1}^{N} a_i G_{ij} b_j \neq \vec{a}^T\vec{b}$]

(not an orthonormal basis ('special case')
$\to$ FOR EVERY BASIS

5. Linear Operator A (today $\mathbb{V} \to \mathbb{V}$)

$y = Ax, \quad A(\lambda\vec{a} + \mu\vec{b}) = \lambda A\vec{a} + \mu A\vec{b}$

with basis $\{\vec{e}_i\}$: $A\vec{e}_j = \sum_{i=1}^{N} A_{ij}\vec{e}_j$

!note $AB\vec{x} \neq BA\vec{x}$

6. Matrices A (today only $N \times N$)

Representation of a linear operator in some basis
$\to$ in basis $\{\vec{e}_i\}$

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_{NN} \end{pmatrix} \qquad \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

Matrix Algebra

- $(A+B)_{ij} = A_{ij} + B_{ij}$
- $(\lambda A)_{ij} = \lambda A_{ij}$
- $(AB)_{ij} = \sum A_{ik} B_{kj}$ (again $AB \neq BA$)

7. some definition

(i) the transpose: $A^T \quad (A^T)_{ij} = A_{ji} \qquad (AB)^T = B^T A^T$

(ii) the determinant: $\det A = |A|$
$\to$ properties
- $|A^T| = |A|$
- $|\lambda A| = \lambda^N |A|$
- $|AB| = |A||B| = |B||A|$

(iii) the inverse: $A^{-1} \quad A^{-1}A = \mathbb{1} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \\ 0 & & 1 \end{pmatrix}$
$\to$ properties
- $(A^T)^{-1} = (A^{-1})^T = A^{-T}$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^{-1})^{-1} = A, \quad \det(A^{-1}) = \frac{1}{\det(A)}$

8. special matrices

(i) orthogonal matrices: $A^T = A^{-1}$
note: $y = Ax \Rightarrow y^Ty = x^T \underbrace{A^TA}_{\mathbb{1}} x = x^Tx$

(ii) symmetric matrices: $A^T = A$

(iii) normal matrices: $A^TA = AA^T$

9. Change of basis

$\vec{x} = \sum_{i=1}^{N} x_i\vec{e}_i$ new basis $\{\vec{e}_j'\}$ with $\vec{e}_j' = \sum_{i=1}^{N} S_{ij}\vec{e}_i$ — some transformation

$\vec{x} = \sum_{j=1}^{N} x_j'\vec{e}_j' = \sum_{j=1}^{N} x_j' \sum_{i=1}^{N} S_{ij}\vec{e}_i \Rightarrow x_i = \sum_{j=1}^{N} S_{ij}x_j'$

$\Leftrightarrow x = Sx'$
$x' = S^{-1}x$ 'transformation of components'

Similarity transformation

$y = Ax \Rightarrow Sy' = ASx' \Rightarrow y' = \underbrace{S^{-1}AS}_{A'} x'$

($y \to Sy'$, $x \to Sx'$)

'map something somewhere, do something and map it back'

10. Eigenvectors & Eigenvalues

$Ax = \lambda x$ — x: Eigenvector, λ: Eigenvalue

notation: x^i is the i-th eigenvector with eigenvalue λ_i

- x^i eigenvector $\Rightarrow \alpha x^i$ is also $\to$ convenient to normalize $x^i \to \frac{x^i}{\|x^i\|}$
- every $N\times N$ matrix has N eigenvectors x^i (might be not l.i.) with eigenvalues λ_i (might not be distinct)

for normal matrices

- if all N eigenvalues are distinct
  $\Rightarrow$ all N eigenvectors are orthogonal
- if some eigenvalues coincide
  $\Rightarrow$ can get N orthogonal eigenvectors by construction
- if $\{x^i\}$ orthonormal $\Rightarrow A = \sum_i \lambda_i x^i (x^i)^T$

for general matrices

- if all eigenvalues distinct $\Rightarrow$ all eigenvectors are l.i.
- if some λ_i coincide $\Rightarrow$ eigenvectors might or might not be l.i.

11. Diagonalisation of matrices

Say A is represented in a basis $\{\vec{e}_i\}$
consider new basis $\vec{x}^j = \sum_{i=1}^{N} S_{ij}\vec{e}_i$ of eigenvectors

$S = \begin{pmatrix} | & & | \\ \vec{x}_1 & \cdots & \vec{x}_N \\ | & & | \end{pmatrix}$

$D = S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}$

note: a $N\times N$ matrix can be diagonalised if N eigenvectors are l.i.

12. Determination of eigenvectors and eigenvalues

$Ax = \lambda x \Leftrightarrow (A - \lambda I)x = \vec{0}$

$|A - \lambda I| \stackrel{!}{=} 0$ characteristic polynomial

- solve $\to N$ eigenvalues λ_i
- plug λ_i into $Ax = \lambda x$
- solve $\to N$ eigenvectors x^i

13. generalised eigenvector problem

$Ax = \lambda x$ and $A = A^T \Rightarrow \overbrace{x^1 .. x^N}^{\text{eig. vec.}}$ basis and $A = SDS^{-1}$ — $\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \\ 0 & & \lambda_N \end{pmatrix}$
and $x^{iT}x^j = 0$ if $\lambda_i \neq \lambda_j$

$Ax = \lambda Bx$ and $A^T = A, B^T = B$ and $x^TBx \geq 0 \quad \forall x$
$\Rightarrow \{x^i\}$ form a basis and $A = BSDS^{-1}$
and $\underbrace{x^{iT}Bx^j = 0}_{\text{"B-orthogonal"}}$ if $\lambda_i \neq \lambda_j$

14. Lagrange multipliers

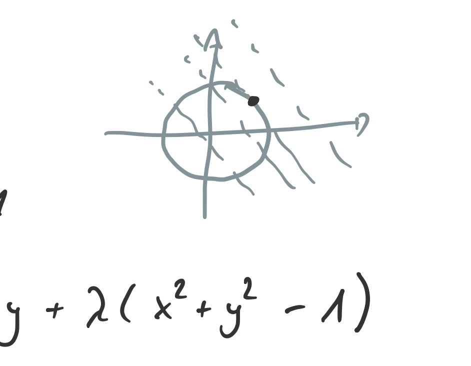
getting started 2 dim., 1 constraint

basic problem: min $f(x,y)$ under constraint $g(x,y) = c$

- naive approach:
  - solve $g(x,y) = c$ for y
  - $f(x, y(x)) \to \frac{\partial f}{\partial x}$

  $\rightsquigarrow$ might be really hard

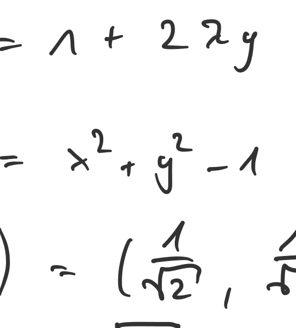
- Lagrange approach:
  want to find a point where the contours touch



$\to$ gradients have to be parallel: $\nabla_{x,y} f = -\lambda \nabla_{x,y} g$

$\Rightarrow$ Lagrange function: $L(x,y,\lambda) := f(x,y) + \lambda \underbrace{(g(x,y) - c)}_{0}$

example & recipe



$f(x,y) = x + y$
$g(x,y) = x^2 + y^2 = 1$

$L(x,y,\lambda) = x + y + \lambda(x^2 + y^2 - 1)$

$\nabla_{x,y,\lambda} \stackrel{!}{=} 0$:

$\frac{\partial L}{\partial x} = 1 + 2\lambda x = 0$
$\frac{\partial L}{\partial y} = 1 + 2\lambda y = 0$
$\Rightarrow x = -\frac{1}{2\lambda} \quad y = -\frac{1}{2\lambda}$

$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - 1 = 0 \to \lambda = \pm\frac{1}{\sqrt{2}}$

$(x^*, y^*) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$

$f(x^*, y^*) = \sqrt{2}$ (max) $\qquad -\sqrt{2}$ (min)

$\{ x \in \mathbb{R}^n \quad L(x,\lambda) = f(x) + \sum_{k=1} \lambda_k g_k$