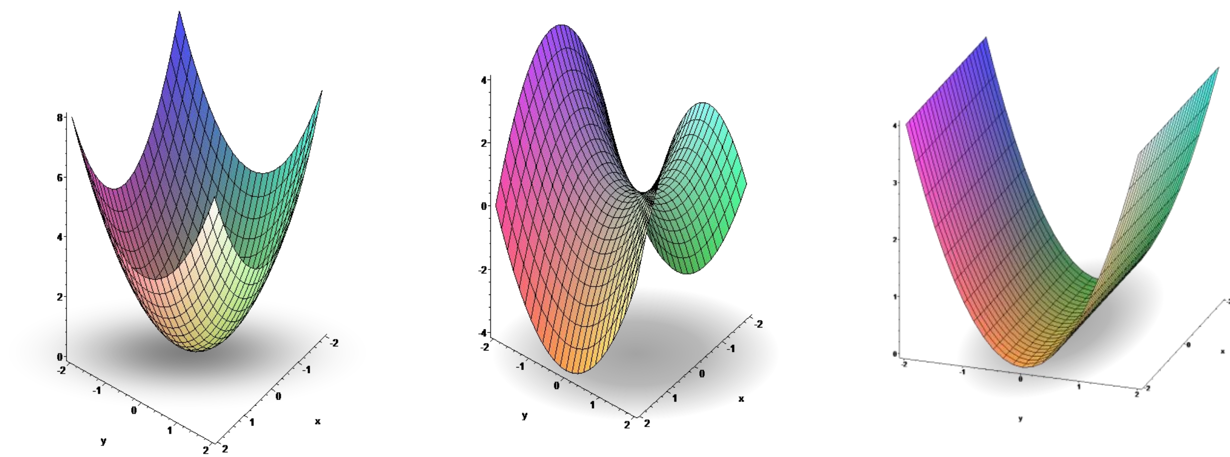
Modelling 1

SUMMER TERM 2020



LECTURE 11

Quadratic Functions

Michael Wand · Institut für Informatik · Michael.Wand@uni-mainz.de

“Non-Linear Linear Algebra”

Quadratic Forms

Multivariate Polynomials

Multi-variate polynomial of total degree $\boldsymbol{d}$

- Polynomial function $f: \mathbb{R}^n \to \mathbb{R}$
- Any 1D section $f(\mathbf{x}_0 + t \cdot \mathbf{r})$ of degree $\leq d$ in t.
- Sum of degrees must be $\leq d$

Examples:

- $f(x, y) := x + xy + y$ has total degree 2. (diagonal!)
- General quadratic polynomial in two variables:

  $f(x, y) \coloneqq c_{20}x^2 + c_{11}xy + c_{02}y^2 + c_{10}x + c_{01}y + c_{00}$

Quadratic Polynomials

General Quadratic Polynomial

- $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \; + \; \mathbf{b}^{\mathrm{T}}\mathbf{x} \; + \; \mathbf{c}$
- $\mathbf{A}$ is an $n \times n$ matrix, $\mathbf{b}$ is an n-dim. vector, $\mathbf{c}$ is a number

Example

$$
\begin{aligned}
f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= [x \; y] \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\
&= [x \; y] \begin{pmatrix} 1x & 2y \\ 3x & 4y \end{pmatrix} \\
&= 1x^2 + (3 + 2)xy + 4y^2 \\
&= 1x^2 + 5xy + 4y^2
\end{aligned}
$$

Quadratic Polynomials

General Quadratic Polynomial

- $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \ + \ \mathbf{b}^{\mathrm{T}}\mathbf{x} \ + \ \mathbf{c}$
- $\mathbf{A}$ is an $n \times n$ matrix, $\mathbf{b}$ is an n-dim. vector, $\mathbf{c}$ is a number

Normalization / Symmetry

- Matrix $\mathbf{A}$ can always be chosen to be symmetric
- If it isn't, we can substitute by $0.5 \cdot (\mathbf{A} + \mathbf{A}^{\mathrm{T}})$, not changing the polynomial

Example

Example:

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = [x\ y]\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= [x\ y]\begin{pmatrix} 1x & 2y \\ 3x & 4y \end{pmatrix}$$

$$= 1x^2 + (3 + 2)xy + 4y^2$$

$$= 1x^2 + (2.5 + 2.5)xy + 4y^2$$

$$= [x\ y]\begin{pmatrix} 1 & 2.5 \\ 2.5 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{x}^T \mathbf{M} \mathbf{x} \\ &= \mathbf{x}^T \frac{1}{2}(\mathbf{M}^T + \mathbf{M})\mathbf{x} \end{aligned}$$

Shapes of Quadrics

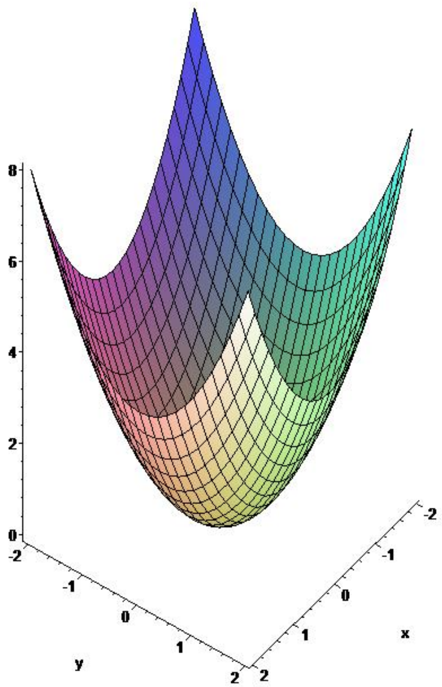
Shape analysis

- $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + \mathbf{c}$
- $\mathbf{A}$ is symmetric
- $\mathbf{A}$ can be diagonalized with orthogonal eigenvectors

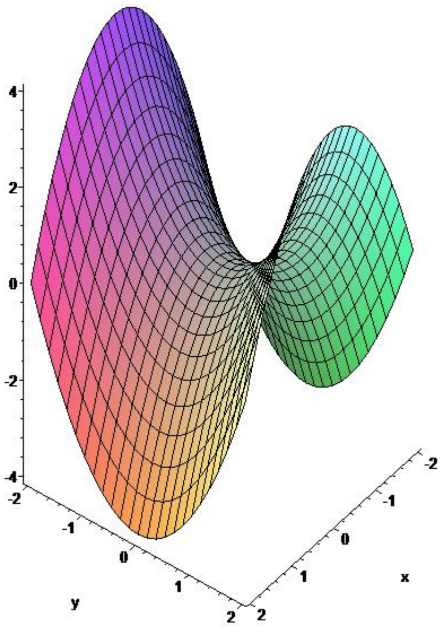
$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}} = \mathbf{U}\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}\mathbf{U}^{\mathrm{T}}$$

  - $\mathbf{U}$ contains principal axes
  - $\mathbf{D}$ gives speeds of growth and up/down direction

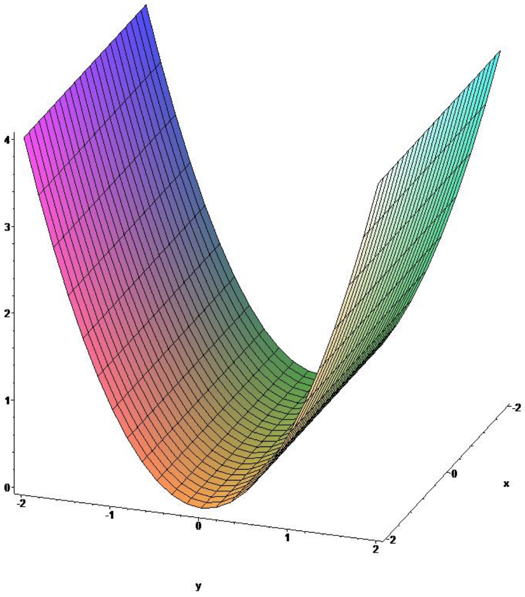
Shapes of Quadratic Polynomials



$\lambda_1 = 1, \lambda_2 = 1$



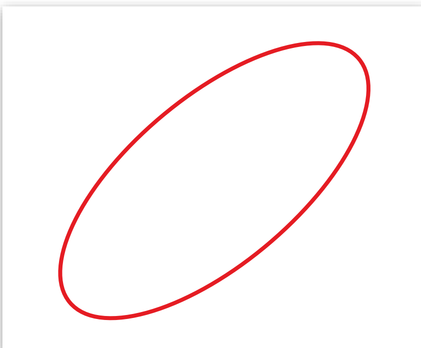
$\lambda_1 = 1, \lambda_2 = -1$



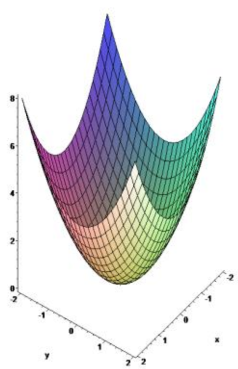
$\lambda_1 = 1, \lambda_2 = 0$

The Iso-Lines: Quadrics

elliptic

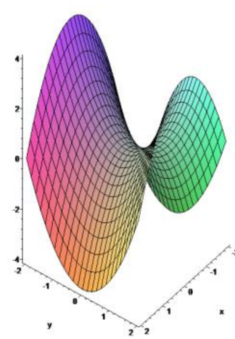


$\lambda_1 > 0, \lambda_2 > 0$

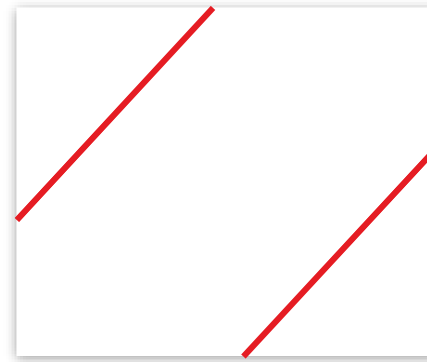


hyperbolic

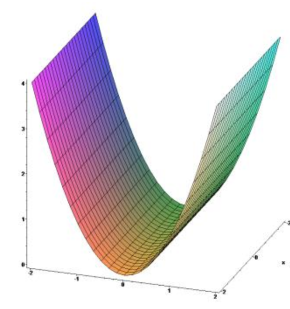
$\lambda_1 < 0, \lambda_2 > 0$



degenerate case



$\lambda_1 = 0, \lambda_2 \neq 0$



“Quadratics”

Quadrics

- Zero level set of a quadratic polynomial: “quadric”
- Shape depends on eigenvalues of $\mathbf{A}$
- $\mathbf{b}$ shifts the object in space
- c sets the level

Quadratic Polynomials

Specifying quadratic polynomials:

- Polynomial: $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} \; + \; \mathbf{c}$
- $\mathbf{b}$: shifts the function in space
  - Assuming full rank $\mathbf{A}$

$$(\mathbf{x} - \boldsymbol{\upmu})^{\mathrm{T}}\mathbf{A}' \, (\mathbf{x} - \boldsymbol{\upmu}) + c'$$

$$= \mathbf{x}^{\mathrm{T}}\mathbf{A}' \, \mathbf{x} - \boldsymbol{\upmu}^{\mathrm{T}}\mathbf{A}' \, \mathbf{x} - \mathbf{x}^{\mathrm{T}}\mathbf{A}'\boldsymbol{\upmu} + \boldsymbol{\upmu}^{\mathrm{T}}\mathbf{A}'\boldsymbol{\upmu} + c'$$

$$= \mathbf{x}^{\mathrm{T}} \underbrace{\mathbf{A}'}_{\mathbf{A}} \, \mathbf{x} + \underbrace{(-2\mathbf{A}' \, \boldsymbol{\upmu})}_{\mathbf{b}} \mathbf{x} + \underbrace{\boldsymbol{\upmu}^{\mathrm{T}}\mathbf{A}'\boldsymbol{\upmu} + c'}_{\mathbf{c}} \quad \rightarrow \quad \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} \; + \; \mathbf{c}$$

- c: additive constant

Some Properties

Important properties

- Multivariate polynomials of fixed degree form a vector space
- We can add them component-wise:

$$
\begin{array}{l}
\quad 2x^2 + 4xy + 3y^2 + 2x + 2y + 4 \\
+\ 3x^2 + 1xy + 2y^2 + 5x + 5y + 5 \\
\hline\hline
=\ 5x^2 + 5xy + 5y^2 + 7x + 7y + 9
\end{array}
$$

Some Properties

Closed Space

- Vector notation:

$$
\begin{aligned}
& \mathbf{x}^{\mathrm{T}}\mathbf{A}_1\mathbf{x} + \mathbf{b}_1{}^{\mathrm{T}}\mathbf{x} + c_1 \\
+ \, & \gamma(\mathbf{x}^{\mathrm{T}}\mathbf{A}_2\mathbf{x} + \mathbf{b}_2{}^{\mathrm{T}}\mathbf{x} + c_2) \\
\hline\hline
= \, & \mathbf{x}^{\mathrm{T}}(\mathbf{A}_1 + \gamma\mathbf{A}_2)\mathbf{x} + (\mathbf{b}_1 + \gamma\mathbf{b}_2)^{\mathrm{T}}\mathbf{x} + (c_1 + \gamma c_2)
\end{aligned}
$$

Quadratic Optimization

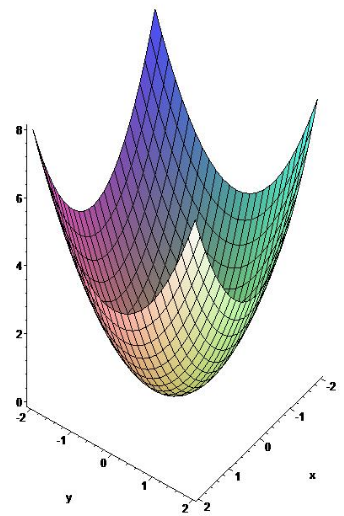
Quadratic Optimization

- Minimize quadratic objective function

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + \mathbf{c}$$

- Required: $\mathbf{A} > 0$ (all eigenvalues positive)
  - It's a paraboloid with a unique minimum
  - The vertex (critical point) can be determined by simply solving a linear system
- Necessary and sufficient condition

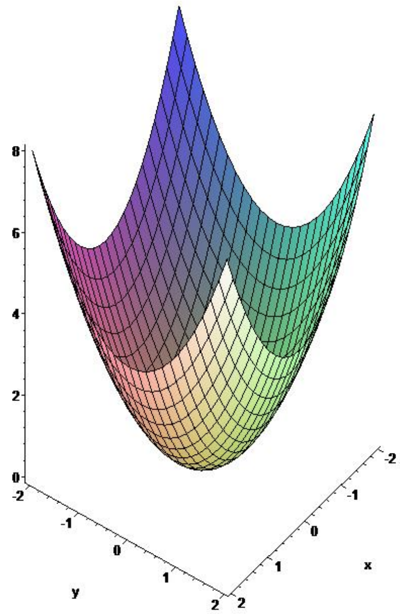
$$2\mathbf{A}\mathbf{x} = -\mathbf{b}$$



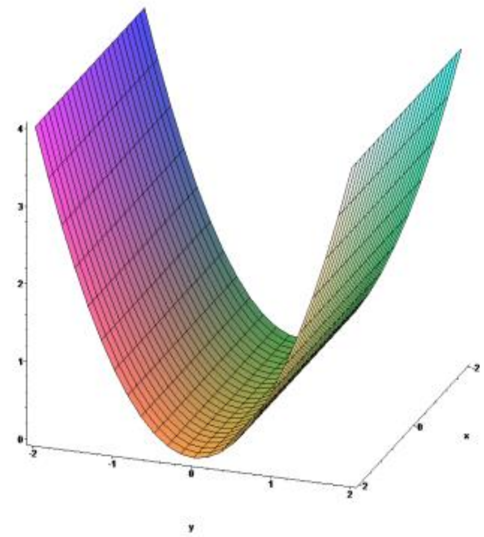
Condition Number

How stable is the solution?

- Depends on Matrix $\boldsymbol{A}$



good



bad

Regularization

Regularization

- Sums of positive semi-definite matrices are positive semi-definite
  - "Valleys can only get steeper"
- Add regularizing quadric
  - "Fill in the valleys"
  - Bias in the solution

Example

- Original: $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + \mathrm{c}$
- Regularized: $\mathbf{x}^{\mathrm{T}}(\mathbf{A} + \epsilon\mathbf{I})\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + \mathrm{c}$

Constraint Optimization

(the other way round)

Rayleigh Quotient

Relation to eigenvalues:

- Min/max eigenvalues of a symmetric $\mathbf{A}$

$$\lambda_{\min} = \min \frac{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}} = \min_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \qquad \lambda_{\max} = \max \frac{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}} = \max_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$$

- The other way round
  - Eigenvalues solve a special constraint optimization problem
  - Hyper-sphere domain
  - “Non-convex”

Coordinate Transformations

One more interesting property:

- Symmetric positive definite ("SPD") matrix $\mathbf{A}$
  - Symmetric
  - All eigenvalues positive
- $\mathbf{A}$ can be written as square of another matrix

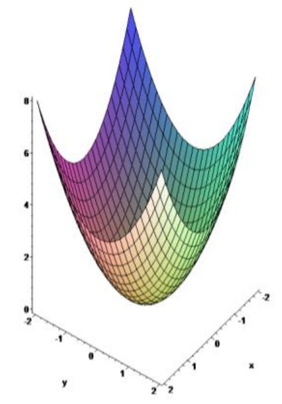
$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}} = \left(\mathbf{U}\sqrt{\boldsymbol{D}}\right) \cdot \left(\sqrt{\boldsymbol{D}}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\right)$$

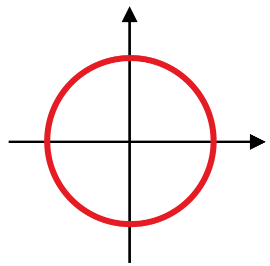
$$\text{"}\sqrt{\mathbf{D}}\text{"} = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_1} \end{pmatrix}$$

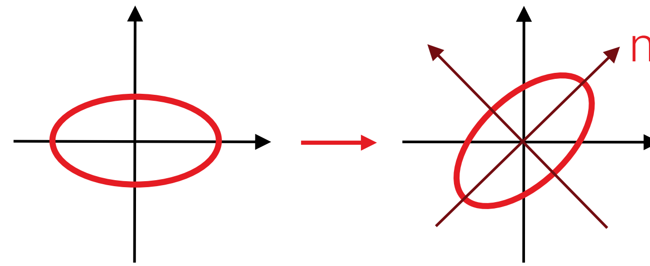
SPD Quadrics

$\mathbf{x}^T\mathbf{x}$

$\mathbf{A} = \mathbf{I}$







$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^T = \left(\mathbf{U}\sqrt{\boldsymbol{D}}\right)^2$$

$$\mathbf{x} \rightarrow \left(\mathbf{U}\sqrt{\boldsymbol{D}}\right)\mathbf{x}$$

Interpretation

- Start with unit quadric $\mathbf{x}^T\mathbf{x}$.
  - Scale the main axis (diagonal of $\mathbf{D}$)
  - Rotate to a different coordinate system (columns of $\mathbf{U}$)
- Recovering main axis from $\mathbf{A}$:
  "principal component analysis" (PCA)