

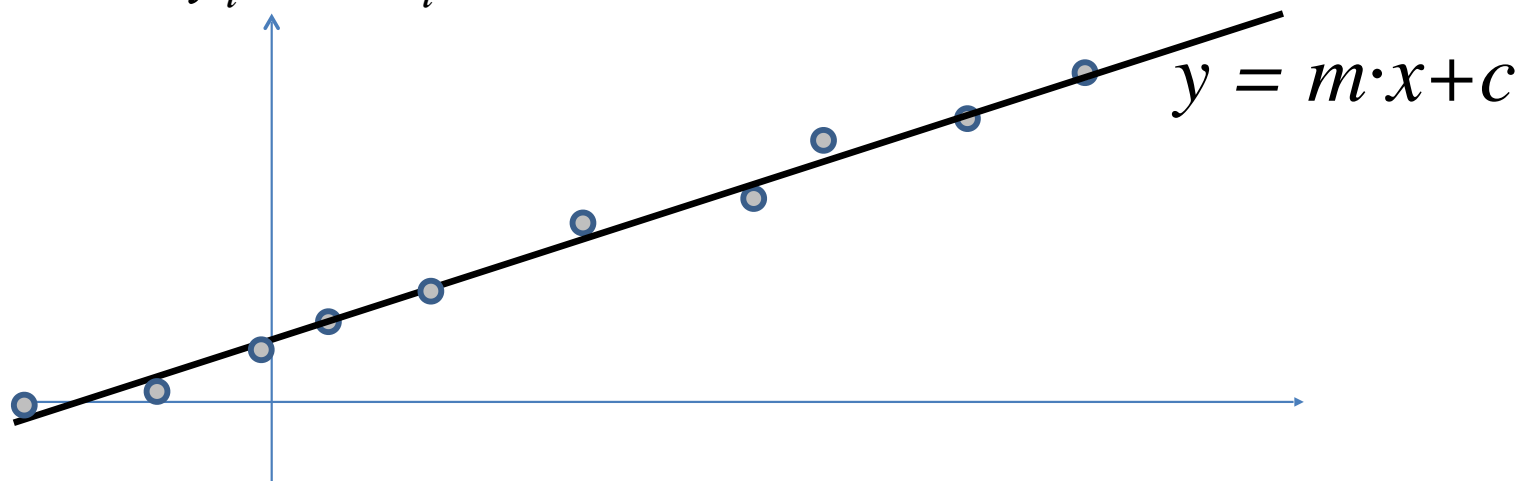
# Algorithmik kontinuierlicher Systeme

## *Lineare Ausgleichsprobleme*



- Gegeben (Mess-)Punkte  
 $[x_1, y_1], [x_2, y_2], [x_3, y_3], \dots, [x_n, y_n]$
- Bestimmen Sie eine Gerade  $y = m \cdot x + c$   
durch diese Punkte  
d.h. bestimmen Sie  $m$  und  $c$  so, dass

$$y_i = m \cdot x_i + c \quad \text{für } i = 1, 2, \dots, n$$



- i.a. nicht exakt lösbar: Suche Näherungslösung

- $A$  sei  $n \times m$  Matrix,  $n > m$ , ( $n$  Zeilen,  $m$  Spalten).
- Ein Gleichungssystem mit mehr Gleichungen als Unbekannten heißt „überbestimmt“.
- Das Gleichungssystem  $Ax = b$  ist dann im Allgemeinen nicht lösbar.
- Selbst wenn die Gleichungen (aus Sicht der Anwendung) eine Lösung haben müssten (z.B. mehrfache Messungen) sind die Gleichungen doch wegen unvermeidlicher Ungenauigkeiten meist widersprüchlich  
=> es ex. keine (exakte) Lösung
- Ausgleichsrechnung / Regressionsanalyse / least squares Approximation

- $Ax = b$  ist nicht lösbar bedeutet:
- das Residuum  $r(x) = b - Ax$  kann durch kein  $x$  zum Verschwinden gebracht werden.
- Es ist nahe liegend, als Ersatz nach einem  $x$  zu suchen, das  $r(x)$  so klein wie möglich macht.
- Man nennt dies das „*lineare Ausgleichsproblem*“
- Methode der „kleinsten Quadrate“ (Quadratsumme):

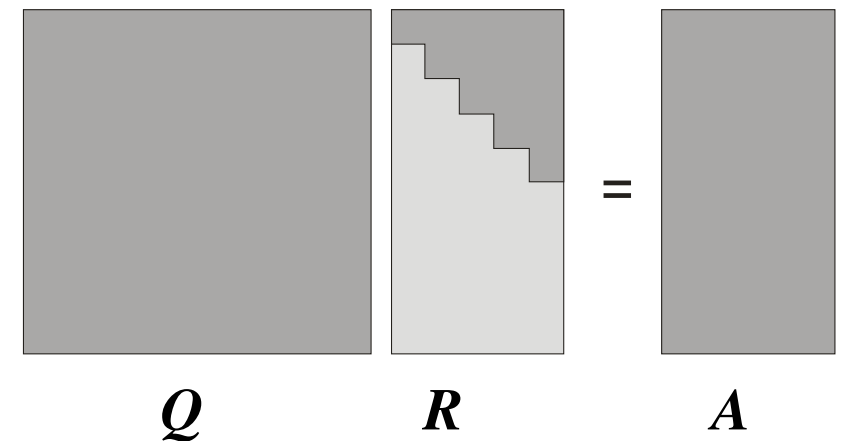
Finde  $\hat{x}$  so dass  $\|r(\hat{x})\| \leq \|r(x)\|$  für alle  $x \in \mathbb{R}^m$

Dabei ist  $\|\cdot\|$  die Euklidische Norm:  $\|r\| = \sqrt{\sum_i r_i^2}$

- Das Ausgleichsproblem hat immer ein eindeutig bestimmtes kleinstes Residuum  $r$ .
- Der dazu gehörige Minimierer  $\hat{x}$  ist nur eindeutig bestimmt, wenn die Spalten von  $A$  linear unabhängig sind, also  $\text{rank}(A) = m$

- Ist die QR-Zerlegung bekannt,

dann gilt

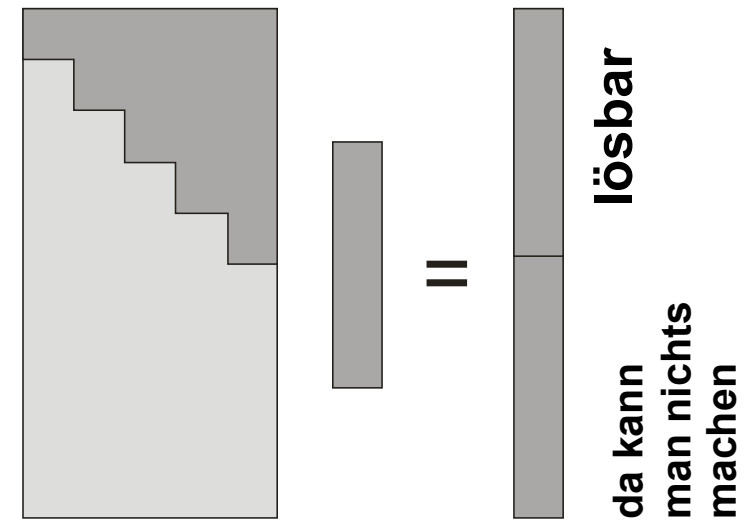


$$\|QRx - b\| = \|Q(Rx - Q^T b)\| = \|Rx - Q^T b\|$$

und somit

$$\|Ax - b\| = \min \Leftrightarrow \|Rx - Q^T b\| = \min$$

- Das Beste, das man erreichen kann, ist die oberen  $m$  Gleichungen zu erfüllen
- Die unteren  $n-m$  „Gleichungen“ spielen keine Rolle.
- Also findet man das beste mögliche  $\hat{x}$ , indem man die oberen  $m$  Gleichungen per Rückwärts-Substitution löst



$$R \quad \hat{x} = \hat{b}$$

lösbar  
da kann man nichts machen

$$(\hat{b} = Q^T b)$$

- Die unteren Komponenten der rechten Seite  $(Q^T b)_j$  mit  $j > m$  bestimmen das Residuum. Sie lassen sich durch keine Wahl von  $x$  zum Verschwinden bringen, sie geben an, wie gut sich  $Ax = b$  überhaupt erfüllen lässt.

- Ein überbestimmtes Gleichungssystem  $Ax = b$  (mit vollem Rang) kann man auch wie folgt lösen:

Bestimme das Minimum des Funktional

$$(x_1, x_2, \dots, x_m) \rightarrow \|b - Ax\|^2 = \sum_{i=1}^n \left( b_i - \sum_{j=1}^m a_{ij} x_j \right)^2$$

- Mathematik: Alle partiellen Ableitungen Null setzen:

$$2 \sum_{i=1}^n (-a_{ik}) \left( b_i - \sum_{j=1}^m a_{ij} x_j \right) = 0 \quad \text{für alle } k = 1, 2, \dots, m.$$

- In Matrix-Schreibweise:  $A^T Ax = A^T b$   
die „Normalengleichung“ des linearen Ausgleichsproblems

- die Normalengleichung des lin. Ausgleichsproblem

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}$$

- Dies ist ein quadratisches LGS, Format  $m \times m$
- $\mathbf{A}^T \mathbf{A}$  ist eine symmetrische  $m \times m$  Matrix
- Falls  $\text{rank}(\mathbf{A}) = m$  dann ist  $\mathbf{A}^T \mathbf{A}$  invertierbar und pos. definit
- Option: Lösen mit Cholesky-Verfahren

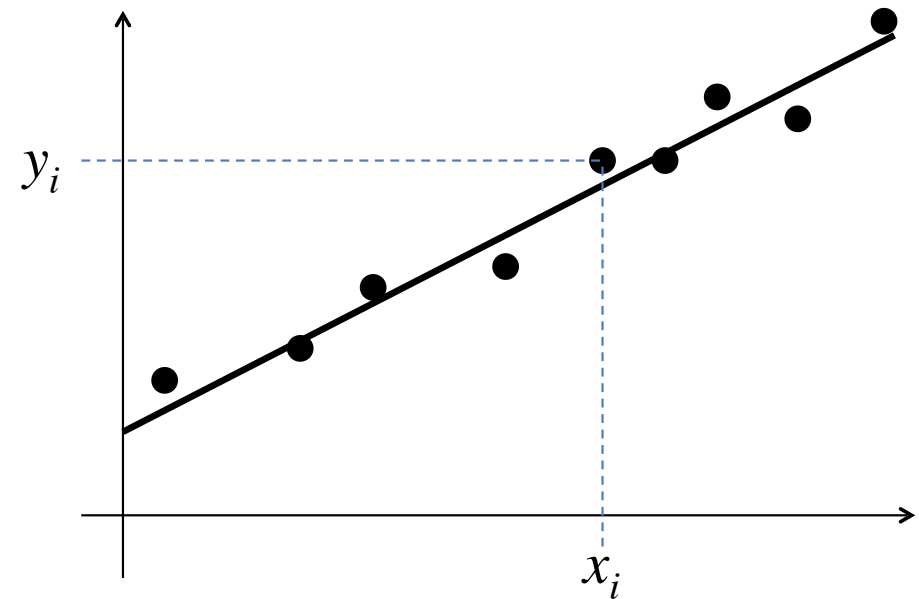
- Gegeben sind eine Reihe von (Meß-)werten

$[x_1, y_1], [x_2, y_2], [x_3, y_3], \dots, [x_n, y_n]$ , (z.B. Strom vs. Spannung)

- Gesucht ist eine Ausgleichsgerade  $y = a_1x + a_0$

- Ideal:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$$



- QR-Zerlegung der Matrix:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \underbrace{\frac{1}{\sqrt{n}} \begin{bmatrix} 1 & \frac{x_1 - \bar{X}}{\sqrt{X^2 - (\bar{X})^2}} \\ 1 & \frac{x_2 - \bar{X}}{\sqrt{X^2 - (\bar{X})^2}} \\ \vdots & \vdots \\ 1 & \frac{x_n - \bar{X}}{\sqrt{X^2 - (\bar{X})^2}} \end{bmatrix}}_Q \underbrace{\sqrt{n} \begin{bmatrix} 1 & \bar{X} \\ 0 & \sqrt{X^2 - (\bar{X})^2} \end{bmatrix}}_R$$

dabei ist  $\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$  und  $\bar{X^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$

Lösung des Ausgleichsproblems:

$$Ra = Q^T Y \quad \text{bzw.} \quad \sqrt{n} \begin{bmatrix} 1 & \bar{X} \\ 0 & \sqrt{\dots - \dots} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \sqrt{n} \begin{bmatrix} \bar{Y} \\ \frac{\overline{XY} - \bar{X} \cdot \bar{Y}}{\sqrt{\dots - \dots}} \end{bmatrix}$$

$$a_1 = \frac{\overline{XY} - \bar{X} \cdot \bar{Y}}{X^2 - (\bar{X})^2}, \quad a_0 = \frac{\bar{Y} \cdot \bar{X^2} - \bar{X} \cdot \overline{XY}}{X^2 - (\bar{X})^2}$$

- Lösen mit Hilfe der Normalengleichung:

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}; A^T A = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} = n \begin{bmatrix} 1 & \bar{X} \\ \bar{X} & \bar{X}^2 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix} = n \begin{bmatrix} \bar{Y} \\ \bar{XY} \end{bmatrix}$$

- Mit Cramer'scher Regel:

$$a_0 = \frac{\bar{Y} \cdot \bar{X}^2 - \bar{X} \cdot \bar{XY}}{\bar{X}^2 - (\bar{X})^2}, \quad a_1 = \frac{\bar{XY} - \bar{X} \cdot \bar{Y}}{\bar{X}^2 - (\bar{X})^2}$$

- Gegeben:
  - $m$  Vektoren  $\mathbf{a}^i$ ,  $1 \leq i \leq m$  der Dimension  $n$ ,  
( $n \gg m$ ), d.h.  $m$  „Muster“.
- Aufgabe:
  - Klassifiziere den Vektor  $\mathbf{b}$  der Dimension  $n$ , d.h. finde heraus, welchem der Muster er am besten entspricht.
- Ein Lösungsansatz:
  - Finde die  $m$  Koeffizienten  $x_i$ ,  $1 \leq i \leq m$ , d.h. den Vektor  $\mathbf{x}$  der Dimension  $m$ , so dass der  $m$ -dimensionale Vektor
    - $\mathbf{y} = a^1 x_1 + a^2 x_2 + \dots + a^m x_m$  die beste Approximation von  $\mathbf{b}$  ist.
  - Erkennung als Muster  $\mathbf{a}^k$ , wenn
    1. der dazugehörige Koeffizient  $x_k$  relativ groß ist *und*
    2. das Residuum  $\|\mathbf{y} - (\mathbf{a}^1 x_1 + \mathbf{a}^2 x_2 + \dots + \mathbf{a}^m x_m)\|$  klein ist
- Entwicklung besserer Verfahren in den Vorlesungen zur *Mustererkennung*



- Bestimmung einer Ausgleichsfunktion (zB. Parabel)

Suche unter allen Parabeln  $y = a_0 + a_1 x + a_2 x^2$  diejenige, die zu den Messwerten

$$[x_1, y_1], [x_2, y_2], [x_3, y_3], \dots, [x_n, y_n],$$

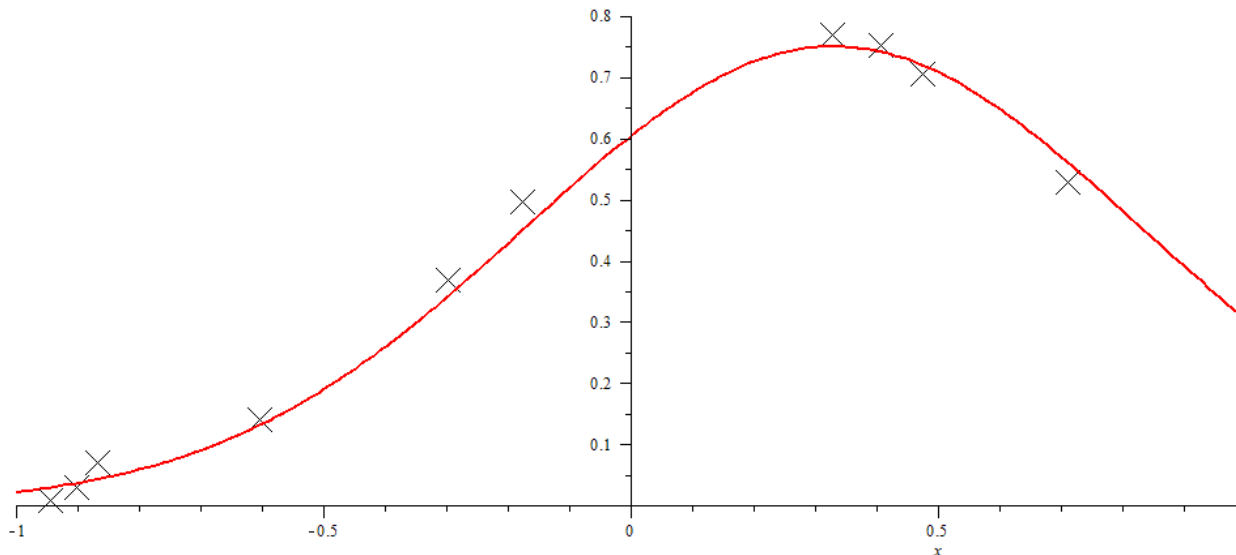
am besten passt.

- Bestimme einer Ausgleichsebene: Gegeben Meß-Daten:
  1. Punkte in der Ebene:  $[x_1, y_1], [x_2, y_2], [x_3, y_3], \dots, [x_n, y_n]$
  2. zu jedem einen Meßwert  $z_1, z_2, z_3, \dots, z_n$

Bestimme diejenige Ebene  $z = a + b x + c y$  die die Meß-Daten am besten approximiert (im quadratischem Mittel)

- Gegeben Messwerte  $(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n),$
- Bestimme eine Gaussfunktion  $f(x, [a, \mu, \sigma]) = a \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$  die die Messwerte bestmöglichst approximiert (im quadratischen Mittel)

$$\sum_{i=1}^n |y_i - f(x_i, [a, \mu, \sigma])|^2 = \min$$



- Erkennung einer Notbremsung aus getakteten Abstandsmessungen
- Gemessene Abstände zum voraus fahrenden Fahrzeug im Takt von 0.1 Sekunden
  - Zeitfenster z.B. 4 Sekunden (d.h. jeweils 40 sukzessive Abstandsmessungen zur Analyse verfügbar)
  - Messungen mit Fehlern
- Methode basierend auf *Least-Square-Fitting*

- Modell:
  - Das eigene Fahrzeug wird als Bezugspunkt angenommen
  - Im normalen Verkehr bewegt sich ein vorausfahrendes Fahrzeug relativ zum eigenen mit konstanter Beschleunigung also einer
  - „Überlagerung“ (mathematisch: Linearkombination) der folgenden „Bewegungs-Grundmuster“
    - Verschwindende Relativgeschwindigkeit:  
gleichbleibender Abstand
    - Konstante Relativgeschwindigkeit:  
linear wachsender/schrumpfender Abstand
    - Konstante Relativbeschleunigung:  
quadratisch wachsender/schrumpfender Abstand

- Wir haben also folgende unbekannte Parameter (für  $m$  Zeitpunkte  $t_0, t_1, t_2, \dots, t_{m-1}$ ) zu bestimmen:

- Position:  $p$
- Geschwindigkeit:  $v$
- Beschleunigung:  $2a$

$$\begin{array}{rclcl}
 i = 0 : & p & + & vt_0 & + & at_0^2 & = & b_0 \\
 i = 1 : & p & + & vt_1 & + & at_1^2 & = & b_1 \\
 & \vdots & & \vdots & & \vdots & & \vdots \\
 i = m - 1 : & p & + & vt_{m-1} & + & at_{m-1}^2 & = & b_{m-1}
 \end{array}$$

- Modell: Abstand  $y$

- $y = p + vt + at^2$

- Für die  $m$  Messwerte erhalten wir  $m$  Gleichungen für 3 unbekannte Parameter

- Berechne  $p, v, a$  als Lösung des linearen Ausgleichsproblems

$$\underbrace{\begin{bmatrix} 1 & t_0 & t_0^2 \\ 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{m-1} & t_{m-1}^2 \end{bmatrix}}_{= A} \begin{bmatrix} p \\ v \\ a \end{bmatrix} = \underbrace{\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{m-1} \end{bmatrix}}_{= b}$$

$$Ax = b$$

- Mit der Lösung als „Least-Squares“-Problem wird im Sinne der 2-Norm die bestmögliche Parameterschätzung gefunden.
- Nimmt man an, dass das Modell korrekt und die Messfehler stochastisch normalverteilt sind, ist dies auch die „wahrscheinlichste“ Wahl.
- Das Modell muss nun noch um das Auftreten einer „Notbremsung“ erweitert werden.
  - Modell:
    - Notbremsung = plötzlich auftretende negative Beschleunigung (d.h. Verzögerung) relativ zum eigenen Fahrzeug:
    - Beginn einer Verzögerung von  $1 \text{ m/s}^2$  vor  $k$  1/10-tel Sekunden mit
      - $z = - (1/10)^2/2 (k^2, (k-1)^2, \dots, 2^2, 1^2, 0, \dots, 0)^T$
      - Bei  $k=4$ , z.B.  $z = -1/100 (16, 9, 4, 1, 0, 0, \dots, 0)^T$
  - Ergänzung der Matrix  $A$  um die Spalte  $z$ .
  - Eine Notbremsung des Vordermanns gilt als erkannt, wenn diese Komponente in der Folge der Abstandsmessungen „signifikant“ auftritt
  - Dies lässt sich an der Größe es dafür errechneten Parameters abschätzen.

- Löse als neues Ausgleichsproblem

$$\tilde{A}\tilde{x} = y, \text{ mit } \tilde{A} = [A|z] \text{ und } \tilde{x} = \begin{bmatrix} p \\ v \\ a \\ s \end{bmatrix}.$$

und bestimme  $s$ .

- Interpretation:  
Die Genauigkeit der Rekonstruktion hängt von der Grösse der (Mess-) Fehler, Länge des Zeitfensters, und vom Zeitintervall stark ab
- In der Praxis würde Python (oder z.B. Matlab o.ä,) eventuell zum Prototyping verwendet werden. Der Algorithmus müsste aber anders implementiert werden um die praktischen Anforderungen zu erfüllen:
  - Realzeit // eingebettetes System // begrenzte Rechenleistung //
  - ....

- „Inkrementelle Verarbeitung“: Ausnutzen, dass in jedem Zeitschritt nur eine „Abstandsmessung hinzu kommt, d.h. nur eine Zeile des Gleichungssystems neu dazu entsteht.
  - Für unser Beispiel kann ggf. die existierende QR-Zerlegung aus dem vorangegangenen Zeittakt wieder verwendet werden: es muss nur eine Diagonale (3 bzw. 4 Rotationen) zusätzlich neu eliminiert werden.
  - In diesem Fall braucht man auch nicht alle Zeilen von  $A$  bzw.  $R$  (und  $y$ ) zu speichern sondern nur jeweils das Dreieck  $R$  und die neu hinzugekommene Zeile.

- Verw. „gewichteter Normen“. Minimiere z.B. bzgl:
  - $\| \mathbf{r} \|^2 := d_1^2 r_1^2 + d_2^2 r_2^2 + \dots + d_n^2 r_n^2$   
 für geeignet gewählte Gewichtungparameter  $d_i$

Die Norm ist also definiert als  $\| \mathbf{r} \| := \| \mathbf{D} \mathbf{r} \|$

Im Fall unserer Anwendung könnte man z.B. weiter zurückliegende Messwerte aus der Vergangenheit weniger stark gewichten.
- Noch allgemeiner kann man für jede positiv definite Matrix  $\mathbf{M}$  eine Norm definieren durch  $\| \mathbf{r} \|^2 := \mathbf{r}^T \mathbf{M} \mathbf{r}$
- Beide Fälle lassen sich auf das normale Ausgleichsproblem zurückführen.

- Überbestimmte Gleichungssysteme
- Näherungslösungen: Minimiere Residuum (im *least-square*-Sinne)
- Lösungsverfahren
  - ▶ QR-Zerlegung
  - ▶ Normalengleichung
  - ▶ später: mittels SVD
- Beispiele:
  - ▶ Ausgleichsgerade, Ausgleichsebene
  - ▶ Mustererkennung, Codierung
  - ▶ Notbremsung