3.8 Methode der kleinsten Quadrate

(Least Squares, Normalgleichung)

Ausgangspunkt: Überbestimmtes System.
Mehr Gleichungen als Unbekannte

$$\mathbf{A}\,\mathbf{x} \stackrel{?}{=} \mathbf{b}$$

Sei A eine m x n – Matrix mit m>n und maximal vollem Rang:
rang(A) = n, d.h. A bildet den $\mathbf{R}^m$ in den ganzen $\mathbf{R}^n$ ab.

Das System A x = b ist dann i.A. nicht lösbar!

Versuche, das Problem so gut wie möglich zu lösen!

Minimiere dazu die Abweichung A x - b in passender Norm!

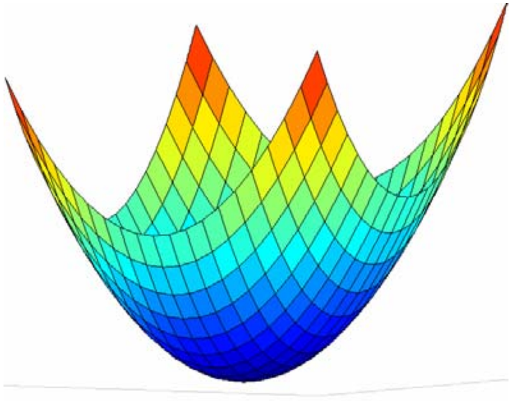
Am besten eignet sich dazu die euklid'sche Norm, da sie auf eine differenzierbare Funktion f führt:

3.8.1. :
$$\min_x \|Ax-b\|_2^2$$

$$f(x_1,\ldots,x_n) := \|Ax-b\|_2^2 =$$

$$= (Ax-b)^T (Ax-b) = x^T A^T Ax - 2x^T A^T b + b^T b =$$

$$= \left\| \begin{pmatrix} (\sum_{j=1}^{n} a_{1,j} x_j) - b_1 \\ \vdots \\ (\sum_{j=1}^{n} a_{m,j} x_j) - b_m \end{pmatrix} \right\|_2^2 = \sum_{k=1}^{m} \left((\sum_{j=1}^{n} a_{k,j} x_j) - b_k \right)^2$$



Die Funktion f beschreibt einen Paraboloiden (n-dim. Parabel). Das eindeutige Minimum dieser Funktion ist an der Stelle, an der die Ableitung gleich Null ist (waagrechte Tangente).

$$0 = \frac{df}{dx_i} = 2\sum_{k=1}^{m}\left(\sum_{j=1}^{n} a_{k,j}x_j - b_k\right)a_{k,i} \qquad \text{für } i=1,\dots,n$$

oder

$$\sum_{k=1}^{m} a_{k,i}\sum_{j=1}^{n} a_{k,j}x_j = \sum_{k=1}^{m} a_{k,i}b_k$$

In Matrixschreibweise: $\left(A^T A x\right)_i = \left(A^T b\right)_i$, i=1,...,n

3.8.2. Normalgleichung zu Ax=b: $A^T A x = A^T b$

Die Matrix **$A^T A$** ist eine n x n – Matrix von Rang n (da A Rang n hat) und beschreibt daher ein eindeutig lösbares, quadratisches lineares Gleichungssystem.

Allerdings ist die Kondition von $A^T A$ oft sehr viel schlechter als die von A, denn:

$$cond_2(A^T A) = \left\| A^T A \right\|_2 \cdot \left\| (A^T A)^{-1} \right\|_2 =$$

$$= \sqrt{\lambda_{\max}(A^T A A^T A) * \lambda_{\max}((A^T A)^{-1}(A^T A)^{-1})} =$$

$$= \lambda_{\max}(A^T A) \cdot \lambda_{\max}(inv(A^T A)) =$$

$$= \lambda_{\max}(A^T A) / \lambda_{\min}(A^T A) = \sigma^2_{\max}(A) / \sigma^2_{\min}(A) =$$

$$= \left\| A \right\|_2^2 \cdot \left\| A^{-1} \right\|_2^2 = cond^2(A)$$

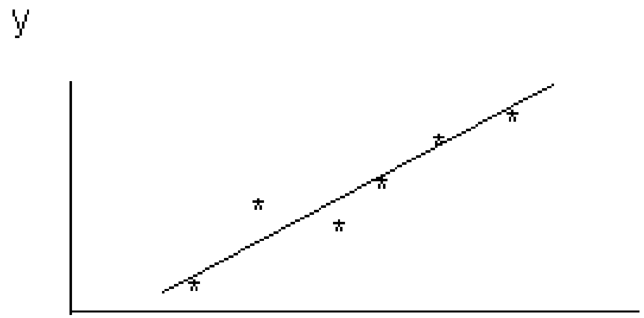
Im folgenden Abschnitt werden wir daher ein besseres Verfahren zur Lösung dieses Problems kennen lernen.

Dazu werden besser orthogonale Matrizen verwendet, um diese Konditionsverschlechterung zu vermeiden.

3.8.3. Lineares Ausgleichsproblem (Ausgleichsgerade)

Gegeben: Punktepaare in der Ebene (x_i, y_i), i=1,...,n;

Gesucht: beste Gerade, die möglichst nahe an den Punkten liegt,



$$\mathbf{y = g(x) = ax + b}\ .$$

Es soll also gelten:

$$\begin{pmatrix} a + bx_1 \\ \vdots \\ a + bx_n \end{pmatrix} \approx \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

oder in Matrixschreibweise

$$A\begin{pmatrix} a \\ b \end{pmatrix} \approx y \quad \text{mit} \quad A = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \quad \text{und} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Die Normalgleichung lautet also

$$A^T A\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} n & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j & \sum_{j=1}^{n} x_j^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} y_j \\ \sum_{j=1}^{n} x_j y_j \end{pmatrix}.$$

Die Lösung dieses 2 x 2 – Gleichungssystems liefert a und b, und damit die gesuchte Gerade y = ax + b .

Allgemeiner:

Ansatzfunktionen $g_1(x),\ldots,g_m(x)$ und

Punkte $(x_1,y_1),\ldots,(x_n,y_n)$, n>m

Gesucht : $f(x)=\sum_{k=1}^{m} a_k g_k(x)$ mit $f(x_j)\approx y_j,\ j=1,\ldots,n$

Mit $G=\begin{pmatrix} g_1(x_1) & \cdots & g_m(x_1) \\ \vdots & & \vdots \\ g_1(x_n) & \cdots & g_m(x_n) \end{pmatrix}$ ist dann $G^T G\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = G^T\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

zu lösen.

Ergebnis ist die ‚näheste' Funktion an den vorgegebenen Punkten, die aus den $g_1,\ldots,g_m$ linear zusammengesetzt ist.

3.9. Die QR-Zerlegung einer Matrix

Schon vorher haben wir bemerkt:
- cond(U) in GE ev. groß, auch bei kleinem cond(A);
- Für A schlecht konditioniert: was ist der Rang von A!
- cond(A^T A) groß,
- aber cond(QA) = cond(A), falls Q orthogonal.

Also sind orthogonale Matrizen sehr gut für äquivalente Umformungen von A geeignet (vgl. LU-Zerlegung).

Außerdem ist $\mathbf{Q^{-1} = Q^T}$ **,** also sind Gleichungssysteme in **Q** leicht zu lösen.

Versuche daher, analog zu A=LR , eine Zerlegung der Form

3.9.1. $\boxed{\mathbf{A = QR}}$

zu finden, mit Q orthogonal und R obere Dreiecksmatrix.

3.9.2 Elementare orthogonale Matrizen

Orthogonale 2 x 2 – Matrix :

$$G = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix}$$ **Givensreflexion.** Denn

$$G^T G = GG = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix} \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix} =$$

$$= \begin{pmatrix} \cos^2(\varphi) + \sin^2(\varphi) & \cos(\varphi)\sin(\varphi) - \sin(\varphi)\cos(\varphi) \\ \sin(\varphi)\cos(\varphi) - \cos(\varphi)\sin(\varphi) & \sin^2(\varphi) + \cos^2(\varphi) \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

Alternativ Givensrotation: $\begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}$

G ist eindeutig bestimmt durch den ‘Winkel’ φ.
Bestimme nun φ so, dass

$$\tilde{A} = GA = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & \tilde{a}_{22} \end{pmatrix}$$ obere Dreiecksmatrix wird.

Dazu muss gelten:

$$\tilde{a}_{21} = \begin{pmatrix} \sin(\varphi) & -\cos(\varphi) \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = \sin(\varphi) a_{11} - \cos(\varphi) a_{21} \overset{!}{=} 0$$

Lösung: $\cot(\varphi) = \dfrac{a_{11}}{a_{21}}; \quad \varphi = arcctg\left(\dfrac{a_{11}}{a_{21}}\right)$ oder $\varphi = arctg\left(\dfrac{a_{21}}{a_{11}}\right)$

Ist $a_{21} = 0$, so ist keine weitere Transformation nötig!

Numerisch stabilere Art der Berechnung :

(a_{21} oder a_{11} könnten fast 0 sein):

$$\rho = sign(a_{11})\sqrt{a_{11}^2 + a_{21}^2}; \qquad \cos(\varphi) = \frac{a_{11}}{\rho}; \qquad \sin(\varphi) = \frac{a_{21}}{\rho};$$

3.9.3. Givens-Reflexion für den allgemeinen n x n – Fall:

Im Wesentlichen Einheitsmatrix, bis auf 2 x 2 – Block, der wie oben definiert ist, abhängig von φ.

Multiplikation mit allgemeiner Givens-Matrix:

$$G_{ij} = \begin{array}{cc} \begin{array}{ccccccccccc} & & & j & & & & i & & & \end{array} & \\ \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & \cos(\varphi) & & & & \sin(\varphi) & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & \sin(\varphi) & & & & -\cos(\varphi) & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix} & \begin{array}{c} \\ \\ \\ j \\ \\ \\ \\ i \\ \\ \\ \\ \end{array} \end{array}$$

zur Elimination eines Elementes a_{ij} der Matrix A .
Dazu multiplizieren wir $G_{ij} \cdot A$.
Dieses Produkt verändert nur die i-te und die j-te Zeile von A.
Es genügt, vom Gesamt-System nur diesen 2 x 2 – Teil zu betrachten. Also muss wieder

$$\varphi = arcctg\left(\frac{a_{jj}}{a_{ij}}\right)$$

gesetzt sein wie oben. (1 → j und 2 → i)

Mit einer solchen Matrix G_{21} wird dann im ersten Schritt a_{21} zu Null gemacht.

$$G_{21} = \begin{pmatrix} G & 0 \\ 0 & I \end{pmatrix}, \quad G = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$

Genauere Analyse eines allgemeinen Eliminationsschritts:

$$\begin{array}{c} \\ j \\ \\ i \\ \\ \end{array} \begin{pmatrix} \ddots & & & \\ & c & & s \\ & & \ddots & \\ & s & & -c \\ & & & \ddots \end{pmatrix} \cdot \begin{pmatrix} \ddots & & & & \\ 0 & \boxed{a_{jj}} & & a_{ji} & \\ 0 & 0 & \ddots & & \\ 0 & \boxed{a_{ij}} & & a_{ii} & \\ \vdots & \vdots & & & \ddots \end{pmatrix} = G_{i,j} \cdot A$$

verändert nur j-te und i-te Zeile;

i-te Zeile: $a_{ik} \to s a_{jk} - c a_{ik}$ für k=j,...,n

speziell: $a_{ij} \to s a_{jj} - c a_{ij} = 0$! soll Null werden und legt daher φ fest.

j-te Zeile: $a_{jk} \to c a_{jk} + s a_{ik}$ für k=j,...,n

mit c und s zu obigen φ .

Verwende der Reihe nach $G_{21}, G_{31}, \ldots, G_{n1}$ zur Bearbeitung der ersten Spalte, also um $a_{21}, a_{31}, \ldots, a_{n1}$ zu Null zu machen,

und danach $G_{32}, G_{42}, \ldots, G_{n2}$, ... , G_{n-1n-2}, G_{nn-2} , und G_{nn-1} um $a_{32}, a_{42}, \ldots, a_{n2}$, ... , a_{n-1n-2}, a_{nn-2} , und a_{nn-1} zu Null zu machen.

Die Reihenfolge, in der die $a_{j,i}$ zu Null gemacht werden, ist gegeben durch:

$$
\begin{array}{ccccc}
\boxdot & & & & \cdot \\
1 & \boxdot & & & \\
2 & n & \boxdot & & \\
\vdots & \vdots & & \boxdot & \\
n-1 & 2n-3 & \cdots & n(n-1)/2 & \boxdot
\end{array}
$$

Jeweils nötig ist eine Multiplikation mit Givens-Reflexion

$G_{i,j}$, i=1,...,n-1 und **j=i+1,...,n.**

Also benötigt man insgesamt n(n-1)/2 Givensreflexionen um eine quadratische n x n –Matrix auf Dreiecksgestalt zu transformieren.

Man benutze also immer das Diagonalelement a_{jj} und eine Kombination von i-ter/j-ter Zeile, um a_{ij} zu Null zu machen.

Mit $$Q^T := G_{n,n-1} G_{n,n-2} \cdots G_{n-1,n-2} \cdots G_{n1} \cdots G_{21}$$

ergibt sich also

$$Q^T A = G_{n,n-1} G_{n,n-2} \cdots G_{n-1,n-2} \cdots G_{n1} \cdots G_{21} A = R$$

mit einer oberen Dreiecksmatrix R.

Dann ist $Q = G_{21} \cdots G_{n1} \cdots G_{n,n-1}$, da $G_{ij}^T = G_{ij}$ und A=Q*R.

Q ist gegeben durch die einzelnen G_{ij};

jedes G_{ij} ist eindeutig gegeben durch das φ_{ij}, das nötig war, um genau ein a_{ij} zu eliminieren.

Genauso kann man für eine **m** x **n** Matrix **A** **(m>n) rank(A)=n** eine QR- Zerlegung berechnen mit

$$\boxed{\mathbf{A}} = \boxed{\quad \mathbf{Q} \quad} \cdot \boxed{\mathbf{R}}$$

Wie bei der Gauss-Elimination eliminiert man also mit den Diagonalelementen der Reihe nach sämtliche Unterdiagonalelemente.

Der Vorteil der QR-Zerlegung:

cond(A) = cond(QR) = cond(R)

Gut für schlecht konditionierte Systeme

Anwendbar auf rechteckige Systeme

Andere Orthogonalisierungsverfahren:

- Gram-Schmidt (orthonormalisiere Vektoren)
- Householder (erzeuge in einem Schritt eine ganze Nullspalte).

3.9.4. Anwendung bei Linearer Ausgleichsrechnung:

$$\min_x \|Ax - b\|_2 = \min_x \|QRx - b\|_2 =$$

$$= \min_x \left\|Q^T (QRx - b)\right\|_2 = \min_x \left\|Rx - Q^T b\right\|_2$$

da Q orthogonal und euklid’sche Norm.

R ist obere Dreiecksmatrix der Dimension m x n und vollen Ranges n.

$$\mathbf{R} = \begin{pmatrix} \mathbf{R_1} \\ \mathbf{0} \end{pmatrix}$$

Das obige Minimum erhält man durch

$$\min_x \left\| \begin{pmatrix} R_1 \\ 0 \end{pmatrix} x - Q^T b \right\|_2^2 = \min_x \left\| \begin{pmatrix} R_1 x \\ 0 \end{pmatrix} - \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{pmatrix} \right\|_2^2 = \min_x \left\| R_1 x - \tilde{b}_1 \right\|_2^2 + \left\| \tilde{b}_2 \right\|_2^2$$

aus der Lösung des Dreieckssystems

$$R_1 x = \tilde{b}_1 .$$

Der Wert des Minimums ist gegeben durch $\left\| \tilde{b}_2 \right\|_2$.

Beispiel: QR-Zerlegung für Least Squares:

$$\min_x \left\| \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} x - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\|_2 = \min_x \|Ax - b\|_2$$

Erster Schritt: a_{21} → 0 :

$$\begin{pmatrix} c & s & 0 \\ s & -c & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} c+s & c+s \\ s-c & s-c \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

mit $\quad c = s = 1/\sqrt{2}, \quad \varphi = \pi / 4$

Zweiter Schritt: a_{32} → 0 :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & s & -c \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & s \\ 0 & -c \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

mit $c = 0, \quad s = 1, \quad \varphi = \pi / 2$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$Q^T \cdot A = R$$

Also

$$Q = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix}$$

Anwendung auf Minimierungsproblem :

$$\min_x \left\| Ax - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\|_2 = \min_x \left\| \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 1 \\ 0 & 0 \end{pmatrix} x - \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\|_2 =$$

$$= \min_x \left\| \begin{pmatrix} \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 1 \end{pmatrix} \cdot x - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \hline 0 \end{pmatrix} \right\|_2 = \min_x \left\| \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 1 \end{pmatrix} \cdot x - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\|_2$$

Lösung x als Lösung des Dreiecksgleichungssystems: $x = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

In diesem Fall liefert x sogar eine genaue Lösung von Ax=b, da der Fehlerterm **$||b_2||$** gleich Null ist.

QR-Zerlegung ist in dieser Form anwendbar für beliebige rechteckige Matrix A, so lange A vollen Rang besitzt.

Kosten des QR-Verfahrens mit Givens für n x n – Matrix:
$2n^3 + O(n^2)$ (also teuerer als Gauss-Elimination mit $2n^3/3$)

Ein Eliminationsschritt bei Spalte k: (2 mult +1 add)2k = 6k flop´s

Insgesamt: $\sum_{k=n-1}^{2} (k-1)*6k = 2n^3 + O(n^2)$

Bei m x n – Matrix mit m>n und Rang(A)=n: $n^2\,(3m-n)$

Anwendung des QR-Verfahren bei

- schlecht konditioniertem Gleichungssystem
- überbestimmtem Gleichungssystem mit vollem Rang (an Stelle der Normalgleichung), wie oben beschrieben
- allgemeinem nichtquadratischen System in der Form QAP = R mit Permutation P zum Vertauschen von Spalten.
(P ist nötig, um einen Block vollen Ranges nach vorne/oben zu transportieren, Beispiel $\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$)

und zur

- **Entdeckung fast linear abhängiger (eigentlich überflüssiger) Gleichungen (numerische Bestimmung des Rangs von A)**
- **Reduktion der Matrix auf den wesentlichen Teil (Noise-reduction)**

3.10 Regularisierung

In vielen praktischen Anwendungen hat man zwar ein überbestimmtes lineares Gleichungssystem vorliegen, aber so, dass die Normalmatrix $\mathbf{A^T A}$ auch noch fast singulär ist!

Dadurch erhält man bei der Lösung dieses Problems einen Vektor x, der sehr groß ist:

Ist in Bx=b die Matrix B fast singulär → ||inv(B)|| sehr groß
→ || x || = || inv(B)*b || sehr groß

Durch Mess/Rundungsfehler enthält aber die rechte Seite b viele kleine Störungen (noise, Rauschen), die in x dann sehr groß werden, so dass die eigentlich genau berechnete Lösung x unbrauchbar ist.

$$\tilde{x} = B^{-1}(b + \Delta b) = B^{-1}b + B^{-1}\Delta b = x + \boxed{B^{-1}\Delta b}$$

Ausweg:

Suche ‚vernünftige' Least Squares Lösung durch Minimierung mit Nebenbedingung: ‚x soll nicht zu groß werden'.

$$\min_x \left(\|Ax-b\|_2^2 + \gamma^2 \|x\|_2^2 \right)$$

Minimerung ←→ Nullstelle der Ableitung führt auf das sog. regularisierte Gleichungssystem

$$\boxed{\left(A^T A + \gamma^2 I\right) x = A^T b}$$

Idee: Verschiebe $A^T A$ durch Aufaddieren von $\gamma^2 I$, so dass die neue Matrix besser konditioniert ist.

Dann ist $\left\| inv\left(A^T A + \gamma^2 I\right) \right\|_2 << \left\| inv(A^T A) \right\|_2$

Daher führen in dem neuen Gleichungssystem die Rausch-komponenten in b nicht mehr zu einem extremen Anwachsen der Lösung x.
Man weiß, dass die gesuchte Lösung x nicht zu groß sein kann, und dies wird durch die Regularisierung gewährleistet.

γ heißt Regularisierungsparameter und die hier beschriebene Methode heißt Tikhonov-Regularisierung.

Regularisierung wird häufig angewendet bei Problemen der Bildverarbeitung (z.B. bei verrauschten, unscharfen Bilder)

Gauss-Elimination, Normalengleichung und QR-Zerlegung, und Regularisierung sind die wichtigsten Werkzeuge für Ax=b.