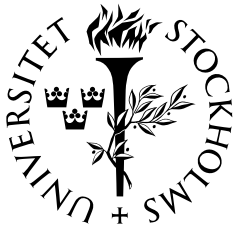


EXAMENSARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

En introduktion till regulariseringsteori för inversa problem

av

Tomas Malm

2008 - No 9

En introduktion till regulariseringsteori för inversa problem

Tomas Malm

Examensarbete i matematik 30 högskolepoäng, fördjupningskurs

Handledare: Hans Rullgård

2008

Abstract

The aim of the thesis is to give an introduction to regularization theory for inverse problems. Though many of the theorems presented are (with small modifications) valid for Hilbert spaces in general, the proofs given here are restricted to euclidean spaces. The methods of Tikhonov regularization and Landweber iteration are established using the concepts of singular value decomposition of a linear operator and regularizing filters. A generalization of Tikhonovs method is also studied. In the last chapter of the thesis we do numerical experiments with a few examples of inverse problems which are common in applications, with the purpose of illustrating the use of regularization methods in numerical computations. For this the problems have to be suitably discretized, since they are originally given in a continuous setting.

En introduktion till regulariseringsteori för inversa problem

Innehåll

1 Introduktion **7**

1.1 Inversa problem och regulariseringsteori 7

1.2 Förberedande linjär algebra 8

2 Regulariseringsmetoder **15**

2.1 Allmän regulariseringsteori och regulariserande filter 15

2.2 Tikhonovregularisering 22

2.3 Generaliserad Tikhonovregularisering 27

2.4 Landweberiteration 32

3 Exempel på inversa problem i tillämpningar; numeriska experiment med regulariseringsmetoder. **36**

3.1 Avfaltningsproblemet i en dimension 36

3.1.1 Numeriska experiment med avfaltningsproblemet i en dimension 41

3.2 Avfaltningsproblemet i två dimensioner 46

3.2.1 Numeriska experiment med avfaltningsproblemet i två dimensioner 49

3.3 Tomografiproblemet och Radontransformen 58

3.3.1 Numeriska experiment med tomografiproblemet 67

Jag vill passa på att rikta ett stort tack till min handledare Hans Rullgård och vill tillägna uppsatsen min mamma och pappa.

1 Introduktion

1.1 Inversa problem och regulariseringsteori

Ett *inverst problem* kan, helt abstrakt, definieras som problemet att för en given avbildning $F : X \longrightarrow Y$ lösa ekvationen

$$F(x) = y,$$

för ett givet $y \in Y$. Det *direkta problemet* består i att för ett givet $x \in X$ bestämma $y = F(x)$. Ett elementärt exempel på ett inverst problem är Lagranges interpolationsproblem, vilket består i att bestämma ett polynom $p(x)$ som antar vissa givna värden $y_1, \ldots, y_n \in \mathbb{R}$ i vissa givna punkter $x_1, \ldots, x_n \in \mathbb{R}$. Motsvarande direkta problem består i att för ett givet polynom $p(x)$ beräkna funktionsvärdena $p(x_1), \ldots, p(x_n)$ i vissa givna punkter $x_1, \ldots, x_n$. I teorin säges ett inverst problem vara *väl ställt* (eng. *well-posed*) om lösningen (1) existerar, (2) är entydigt bestämd och (3) är *stabil*, vilket formellt definieras som att lösningen x beror kontinuerligt på indata y. Annars säges problemet vara *illa ställt* (eng. *ill-posed*). Vad beträffar stabiliteten är det för numeriska beräkningar inte alltid nog att lösningen är stabil i betydelsen kontinuerligt beroende av indata. Lösningen kan också vara instabil i den mindre precist definierade bemärkelsen att den är "mycket känslig för" perturbationer i y, så att små diskrepanser i indata resulterar i stora beräkningsfel. När vi säger att lösningen är *instabil* menar vi här instabil i någon av dessa två bemärkelser.

Många i tillämpningar förekommande inversa problem är *linjära*, vilket innebär att de kan formuleras som en linjär operatorekvation (ett ekvationssystem)

$$Kx = y, \tag{1}$$

där $K : X \longrightarrow Y$ är en linjär avbildning mellan två linjära rum X och Y, typiskt två Hilbertrum. De inversa problem vi intresserar oss för här är linjära.

Låt $K : X \longrightarrow Y$ vara en linjär avbildning mellan de euklidiska rummen X och Y. Från elementär linjär algebra vet vi att rummet X kan skrivas som den direkta summan av nollrummet till K och det ortogonala komplementet till nollrummet:

$$X = N(K) \oplus N(K)^{\perp}. \tag{2}$$

Det följer att om ekvationen (1) har åtminstone en lösning, för ett givet $y \in Y$, så finns en entydigt bestämd lösning $x^* \in N(K)^{\perp}$ till ekvationen och, vidare, att alla andra lösningar erhålles genom superposition av elementen i $N(K)$.

Vi kan alltså inskränka sökandet efter en eventuell lösning till det inversa problemet till det ortogonala komplementet. Om K är injektiv, så är förstås

$N(K)^\perp = X$. Ifall K inte är injektiv, kan vi istället betrakta restriktionen

$$K := K|N(K)^\perp,$$

vilket utgör en injektiv avbildning. Låt oss förenkla det fortsatta resonemanget genom att göra antagandet att K är injektiv på hela X. Låt $K^{-1} : K(X) \subseteq Y \longrightarrow X$ beteckna inversen till K.

Antag nu att (1) har åtminstone en lösning, alltså att högerledet $y \in K(X)$. I tillämpningar, i numeriska beräkningar eller vetenskapliga experiment, kan vi inte utgå ifrån att vi har tillgång till ett "exakt värde" på högerledet y, utan endast ett mätvärde med någon viss felgräns eller felmarginal $\delta > 0$. Låt oss alltså anta att det givna är ett $\delta > 0$ och mätdatan y^δ med

$$\|y - y^\delta\| \leq \delta \tag{3}$$

Detta gör oss hänvisade till att istället försöka lösa den *perturberade* ekvationen

$$Kx^\delta = y^\delta. \tag{4}$$

Ofta låter vi x och y stå för den exakta lösningen till det inversa problemet respektive det exakta högerledet, och x^δ och y^δ för deras respektive perturberade motsvarigheter.

Ett första problem vi kan stöta på är att (4) saknar lösning, alltså att $y^\delta \notin K(X)$. Ett andra problem vi kan stöta på är instabilitet, alltså att lösningen till det inversa problemet är känslig för perturbationer i högerledet, så pass att lösningen $x^\delta = K^{-1}y^\delta$ till oigenkännlighet avviker från den egentliga eller exakta lösningen x.

Det vi i båda dessa fall dock kunde hoppas på är att åtminstone kunna bestämma en approximation $x^\delta \in X$ till den exakta lösningen x sådan att $\|Kx^\delta - y^\delta\|$ inte är avsevärt mycket större än δ. Ett sätt att göra det vore att försöka bestämma en kontinuerlig och mer stabil avbildning

$$R : Y \longrightarrow X$$

som får fungera som en approximation eller utvidgning av inversen K^{-1}. Att på ett metodiskt sätt bestämma en sådan approximerande avbildning är uppgiften för *regulariseringsteorin*, som vi återkommer till i avsnitt 3.1.

1.2 Förberedande linjär algebra

För de resultat i det här avsnittet som inte ges bevis hänvisas läsaren till någon lärobok i linjär algebra, exempelvis [3]. Här förutsättes att de linjära rum X, Y, $Z, \ldots$ vi studerar är *euklidiska rum*, det vill säga ändligtdimensionella linjära rum över $\mathbb{R}$ med skalärprodukt $\langle \cdot, \cdot \rangle$, och alltså isomorfa med $\mathbb{R}^n$ tillsamans med standardskalärprodukten. För att undvika tröttsamma upprepningar kommer vi att låta det vara underförstått i satser

och definitioner, om inget annat säges. Vi låter det också vara underförstått att rummen är normerade, med den av skalärprodukten naturligt inducerade normen

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

I detta avsnitt betecknar $A,\ B,\ C, \ldots$ linjära avbildningar; beteckningen K har vi reserverat för den till ett inverst problem hörande avbildningen. Vi låter $[A] = [A]_{e,f}$ beteckna matrisen för avbildningen A med avseende på baserna e och f för X respektive Y. Vidare låter vi $[x] = [x]_e \in \mathbb{R}^n$ stå för koordinatvektorn (på kolonnform) till ett givet $x \in X$ med avseende på basen e.

I *standardbasen* $e = (e^1, \ldots, e^m)$ för $\mathbb{R}^n$ ges det i:te elementet i e^λ, $1 \leq i, \lambda \leq n$, av

$$e_i^\lambda = \delta_{i\lambda},$$

där funktionen i högerledet är *Kroneckers delta*

$$\delta_{ab} = \begin{cases} 1 & \text{om } a = b, \\ 0 & \text{om } a \neq b. \end{cases}$$

En linjär avbildning A på de normerade rummen X och Y är *begränsad* om det existerar $c > 0$ sådant att

$$\|Ax\| \leq c\|x\|,$$

för alla $x \in X$. I euklidiska rum är alla linjära avbildningar begränsade och kontinuerliga, vilket för Hilbertrum i allmänhet kan visas vara ekvivalenta villkor. Dock är i det oändligt-dimensionella fallet inte alla linjära avbildningar begränsade eller kontinuerliga. För givna normer på X och Y definieras *normen av den linjära avbildningen* A som

$$\|A\| := \sup\{\|Ax\| : \|x\| \leq 1\}.$$

Man visar lätt att

$$\begin{aligned} \sup\{\|Ax\| : \|x\| \leq 1\} &= \sup\left\{ \frac{\|Ax\|}{\|x\|} : x \neq 0 \right\} \\ &= \inf\{c > 0 : \|Ax\| \leq c\|x\|, \text{ för alla } x \in X\}. \end{aligned}$$

Följande räknelagar gäller för operatornormen:

(i) $\|A\| \geq 0$, med likhet endast för nolloperatorn $A \equiv 0$.

(ii) $\|\lambda A\| = |\lambda|\|A\|$, för alla $\lambda \in \mathbb{R}$.

(iii) $\|A + B\| \leq \|A\| + \|B\|$, där $A, B : X \longrightarrow Y$. (Triangelolikheten)

(iv) $\|Ax\| \leq \|A\|\|x\|$, för alla $x \in X$.

(v) $\|BA\| \leq \|B\|\|A\|$, där $A : X \longrightarrow Y$ och $B : Y \longrightarrow Z$.

(vi) $\|A^*\| = \|A\|$.

(vii) $\|A^*A\| = \|A\|^2$.

I de sista två punkterna är A^* den till A adjunkta avbildningen, som allmänt definieras för kontinuerliga linjära avbildningar på Hilbertrum. Denna är en sorts generalisering av matristransponaten kan man säga. I fallet med euklidiska rum gäller nämligen att

$$[A^*] = [A]^t, \tag{5}$$

i givna ON-baser för X och Y, vilket framgår av beviset för följande sats:

Sats 1.2.1. *Låt $A : X \longrightarrow Y$ vara linjär och låt $y \in Y$ vara fixt. Då existerar ett entydigt bestämt $z \in X$ sådant att*

$$\langle Ax, y\rangle = \langle x, z\rangle, \text{ för alla } x \in X.$$

Bevis. Vi börjar med att visa entydigheten. Antag att $z, z' \in X$ uppfyller villkoret. Då har vi att $\langle x, z\rangle = \langle x, z'\rangle$ och alltså $\langle x, z - z'\rangle = 0$, för alla $x \in X$. Alltså är $\langle z - z', z - z'\rangle = 0$, varför $z - z' = 0$. Existensen av z får vi enkelt på följande vis: låt $e = (e_1, \ldots, e_m)$ vara en ON-bas för X och $f = (f_1, \ldots, f_n)$ en ON-bas för Y. För givet $y \in Y$ definierar vi $z \in X$ enligt

$$[z] := [A]^t[y].$$

Vi har då att $\langle Ax, y\rangle = \langle x, z\rangle$ omm $\langle [A][x], [y]\rangle = \langle [x], [z]\rangle$ omm

$$\sum_{i=1}^{n} \Big(\sum_{j=1}^{m} a_{ij}x_j \Big) y_i = \sum_{i=1}^{m} x_i \Big(\sum_{j=1}^{n} a_{ji}y_j \Big).$$

Den sista likheten följer av att både $a_{ij}x_jy_i$ och $a_{ji}x_iy_j$ ingår som termer i vänster- och högerled, för alla $i \in \{1, \ldots, m\}$ och $j \in \{1, \ldots, n\}$. Det vill säga, alla möjliga kombinationer ingår i båda summorna. □

Av satsen följer att A^* är väldefinierad:

Definition 1.2.1. *För en given linjär avbildning $A : X \longrightarrow Y$ definieras* adjunkten till A *eller den till A* adjunkta avbildningen*, $A^* : Y \longrightarrow X$, enligt följande villkor: för ett givet $y \in Y$ är A^*y sådant att*

$$\langle Ax, y\rangle = \langle x, A^*y\rangle, \text{ för alla } x \in X.$$

Sats 1.2.2. *Följande räknelagar gäller för adjunkten:*

(i) $(A^*)^* = A$.

(ii) $(\lambda A)^* = \lambda A^*$, *för* $\lambda \in \mathbb{R}$.

(iii) $(A+B)^* = A^* + B^*$, *där* $A, B : X \longrightarrow Y$.

(iv) $(BA)^* = A^* B^*$, *där* $A : X \longrightarrow Y$ *och* $B : Y \longrightarrow Z$.

(v) Om A *är inverterbar, så är även* A^* *inverterbar och* $(A^*)^{-1} = (A^{-1})^*$.

Bevis. I fallet med euklidiska rum korresponderar dessa lagar uppenbarligen mot räknelagarna för matristransponaten, och är med användande av likheten (5) en omedelbar konsekvens av dem. □

Definition 1.2.2. *En linjär operator* A *på rummet* X *säges vara* symmetrisk *eller* själv-adjungerad *om* $A^* = A$.

Sats 1.2.3 (**Spektralsatsen**)**.** *Om operatorn* A *på* X *är symmetrisk, så existerar en ON-bas för* X *bestående av egenvektorer till* A. *Egenvektorer hörande till skilda egenrum (egenvärden) är inbördes ortogonala.*

Spektralsatsen kan utvidgas till en likartad sats för linjära avbildningar generellt, satsen om singulärvärdesuppdelning, varur spektralsatsen erhålles som ett specialfall. För en given linjär avbildning $A : X \longrightarrow Y$ är sammansättningen $A^*A : X \longrightarrow X$ (och $AA^* : Y \longrightarrow Y$) en symmetrisk operator på X (respektive Y). Detta följer av räknelagarna för adjunkten:

$$(A^*A)^* = A^*(A^*)^* = A^*A.$$

Dessutom gäller att

Sats 1.2.4. *Alla egenvärden till* A^*A *och* AA^* *är icke-negativa och deras respektive positiva egenvärden sammanfaller.*

Bevis. Låt λ vara ett egenvärde till A^*A. Då är

$$\lambda\langle x, x\rangle = \langle A^*Ax, x\rangle = \langle Ax, Ax\rangle \geq 0,$$

för alla $x \in X$. Alltså är $\lambda \geq 0$.

Antag nu att $\lambda > 0$. Om vi kan visa att λ även är ett egenvärde till AA^* så är vi klara, då den omvända implikationen följer av symmetriskäl. Låt $x_0 \neq 0$ vara en egenvektor till A^*A för egenvärdet λ.

Sätt $y_0 = Ax_0$. För det första är $y_0 \neq 0$, eftersom vi annars skulle ha $0 = A^*y_0 = A^*Ax_0 = \lambda x_0$. För det andra har vi att

$$AA^*y_0 = AA^*(Ax_0) = A(\lambda x_0) = \lambda y_0.$$

Alltså är y_0 en egenvektor till AA^* för egenvärdet λ. □

Definition 1.2.3. *Låt* $\lambda_i, i \in I$ *(*I *ändlig), vara egenvärdena till* A^*A. *Talen* $\mu_i = \sqrt{\lambda_i}$ *kallas för* singulärvärdena *till avbildningen* A.

Vi påminner oss om att *multipliciteten* för ett egenvärde är lika med dimensionen av dess motsvarande egenrum. Av sats 1.2.4 följer att A och A^* har samma positiva singulärvärden. Dessutom sammanfaller multipliciteterna för dessa. Dock kan multipliciteten skilja sig åt för singulärvärdet 0, och gör det i allmänhet, såvida inte X och Y är av samma dimension. Vidare kan också den ena av operatorerna A^*A och AA^* ha egenvärdet $\lambda = 0$, medan den andra operatorn inte har det, så att den ena av avbildningarna A respektive A^* har singulärvärdet $\mu = 0$, medan den andra inte har det.

Observera också att

$$0 \leq \mu \leq \|A\|$$

för varje singulärvärde μ till A. Låt nämligen $x \in X$ vara en egenvektor till A^*A för egenvärdet $\lambda = \mu^2$. Låt oss anta att $\|x\| = 1$. Vi har då enligt räknelagarna för operatornormen att

$$\lambda = \|\lambda x\| = \|A^*Ax\| \leq \|A^*A\|\|x\| = \|A\|^2.$$

Sats 1.2.5 (**Singulärvärdesuppdelning**). *Låt $(\mu_j)_{j\in J}$ (J ändlig), med $\mu_1 \geq \mu_2 \geq \mu_3 \geq \ldots > 0$ vara den ordnade följden av positiva singulärvärden till A, räknade med multiplicitet. Det existerar ON-system $(x_j)_{j\in J} \subset X$ och $(y_j)_{j\in J} \subset Y$ sådana att*

$$Ax_j = \mu_j y_j \text{ och } A^*y_j = \mu_j x_j.$$

Den ordnade trippeln $(\mu_j, x_j, y_j)_{j\in J}$ kallas för ett singulärvärdessystem *för A. Varje $x \in X$ har den entydigt bestämda så kallade* singulärvärdesuppdelningen

$$x = x_0 + \sum_{j\in J} \langle x, x_j\rangle x_j,$$

där $x_0 \in N(A)$ och $x - x_0 \in N(A)^\perp$. Vidare är

$$Ax = \sum_{j\in J} \mu_j \langle x, x_j\rangle y_j.$$

Bevis. Enligt spektralsatsen finns det en ON-bas för X bestående av egenvektorer till A^*A. Låt $(\lambda_i)_{i\in I}$ (I ändlig), med $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq 0$ vara den ordnade följden av (som sagt icke-negativa) egenvärden till A^*A, räknade med multiplicitet, och $(e_i)_{i\in I}$ en korresponderande bas av egenvektorer till A^*A. Låt $(x_j)_{j\in J}$ ($|J| \leq |I|$) vara den ordnade följden av de egenvektorer i (e_i) som svarar mot positiva egenvärden. Sätt vidare $y_j = 1/\sqrt{\lambda_j} Ax_j$. Då har vi att

$$\begin{aligned} &Ax_j = \sqrt{\lambda_j} y_j = \mu_j y_j, \quad \text{och} \\ &A^*y_j = A^*(1/\sqrt{\lambda_j} Ax_j) = 1/\sqrt{\lambda_j} A^*Ax_j = \sqrt{\lambda_j} x_j = \mu_j x_j. \end{aligned}$$

Vi ska visa att $(y_j)_{j\in J}$ är ett ON-system i Y. Låt $i, j \in J$. Då har vi att

$$\begin{aligned}\langle y_i, y_j\rangle &= \langle 1/\sqrt{\lambda_j}Ax_i, 1/\sqrt{\lambda_j}Ax_j\rangle = 1/\sqrt{\lambda_i\lambda_j}\langle Ax_i, Ax_j\rangle \\ &= 1/\sqrt{\lambda_i\lambda_j}\langle A^*Ax_i, x_j\rangle = \lambda_i/\sqrt{\lambda_i\lambda_j}\langle x_i, x_j\rangle \\ &= \delta_{ij}\end{aligned}$$

Vidare kan varje $x \in X$ skrivas på formen

$$x = \sum_{i\notin J}\langle x, e_i\rangle e_i + \sum_{j\in J}\langle x, x_j\rangle x_j$$

Sätter vi

$$x_0 = \sum_{i\notin J}\langle x, e_i\rangle e_i,$$

får vi singulärvärdesuppdelningen av x med $x_0 \in N(A)$. Att $x - x_0 \in N(A)^\perp$ följer av att $\langle x_j, x'\rangle = 0$ för alla $j \in J$ och $x' \in N(A)$:

$$\langle x_j, x'\rangle = 1/\mu_j\langle \mu_j x_j, x'\rangle = 1/\mu_j\langle A^*y_j, x'\rangle = 1/\mu_j\langle y_j, Ax'\rangle = 0. \qquad \square$$

Sats 1.2.6. *Låt $(\mu_j, x_j, y_j)_{j\in J}$ vara ett singulärvärdessystem för $A : X \longrightarrow Y$ och låt $y \in Y$ vara godtyckligt. Då är $(\mu_j, y_j, x_j)_{j\in J}$ ett singulärvärdessystem för $A^* : Y \longrightarrow X$, så att alltså:*

(i)

$$A^*y = \sum_{j\in J}\mu_j\langle y, y_j\rangle x_j.$$

(ii) y entydigt kan skrivas på formen

$$y = y_0 + \sum_{j\in J}\langle y, y_j\rangle y_j,$$

där $y_0 \in N(A^*)$ och $y - y_0 \in N(A^*)^\perp$. Vidare gäller att:

(iii) Operatorekvationen

$$Ax = y$$

är lösbar, d.v.s. $y \in A(X)$, om och endast om $y \in N(A^*)^\perp$. I detta fall är

$$x = \sum_{j\in J}\frac{1}{\mu_j}\langle y, y_j\rangle x_j \tag{6}$$

en lösning till ekvationen.

Bevis. (i): Det gäller att $A^*y = z$ omm $\langle Ax, y\rangle = \langle x, z\rangle$, för alla $x \in X$. Vi har att

$$\langle Ax, y\rangle = \Big\langle \sum_{j\in J} \mu_j \langle x, x_j\rangle y_j, y \Big\rangle = \sum_{j\in J} \mu_j \langle x, x_j\rangle \langle y, y_j\rangle, \quad \text{och}$$
$$\langle x, \sum_{j\in J} \mu_j \langle y, y_j\rangle x_j\rangle = \sum_{j\in J} \langle x, \mu_j \langle y, y_j\rangle x_j\rangle = \sum_{j\in J} \mu_j \langle x, x_j\rangle \langle y, y_j\rangle.$$

Alltså gäller likheten i (i).

(ii): Låt $y_0 \in Y$ ges av villkoret att

$$y = y_0 + \sum_{j\in J} \langle y, y_j\rangle y_j.$$

Applicerar vi A^* på vänster- och högerled får vi att

$$A^*y = A^*y_0 + \sum_{j\in J} \mu_j \langle y, y_j\rangle x_j = A^*y_0 + A^*y,$$

varav följer att $y_0 \in N(A^*)$. Att $y - y_0 \in N(A^*)^\perp$ följer av att $\langle y_j, y'\rangle = 0$ för alla $j \in J$ och $y' \in N(A)$, vilket visas på samma sätt som i föregående sats.

(iii): För den första implikationen, antag att x är en lösning till operatorekvationen. Låt $y_0 \in N(A^*)$ vara godtyckligt. Då är

$$\langle y, y_0\rangle = \langle Ax, y_0\rangle = \langle x, A^*y_0\rangle = 0,$$

och alltså ligger y i det ortogonala komplementet.

Omvänt, antag att $y \in N(A^*)^\perp$. Låt x ges av ekvation (6). Vi får att

$$Ax = A\Big(\sum_{j\in J} \frac{1}{\mu_j} \langle y, y_j\rangle x_j\Big) = \sum_{j\in J} \langle y, y_j\rangle y_j = y - y_0,$$

för något $y_0 \in N(A^*)$. Av att $y - y_0$ ligger i det ortogonala komplementet och antagandet att även y gör det följer att $y_0 = 0$. Alltså är x en lösning till operatorekvationen. □

2 Regulariseringsmetoder

2.1 Allmän regulariseringsteori och regulariserande filter

Låt oss nu fortsätta studiet av regulariseringsmetoder där vi slutade och utveckla teorin mer formellt och i detalj. Satser och bevis i detta avsnitt och avsnitten om Tikhonovregularisering och Landweberiteration är väsentligen hämtade ur kapitel 2 av [2]. Avsnittet om generaliserad Tikhonovregularisering använder framställningen i [4]. Vi låter som förut $K : X \longrightarrow Y$ vara en injektiv linjär avbildning på euklidiska rum.

Definition 2.1.1. *En* regulariseringsstrategi (för K) *är en familj av kontinuerliga avbildningar*

$$R_\alpha : Y \longrightarrow X, \quad \alpha \in I \subseteq \mathbb{R}^+$$

sådan att

$$R_\alpha K x \to x, \ \text{då } \alpha \to 0,$$

för alla $x \in X$. (Med andra ord konvergerar avbildningarna $R_\alpha K$ punktvis mot identitetsoperatorn på X.) Vi säger att regulariseringsstrategin är linjär *om R_α är linjär, för varje $\alpha \in I$.*

För ett givet val av α uppfattar vi

$$x^{\alpha,\delta} := R_\alpha y^\delta \tag{7}$$

som en approximation till den exakta lösningen x.

Vi kommer här enbart ägna oss åt linjära regulariseringsstrategier. För detta fall följer kontinuiteten av att de regulariserande avbildningarna är linjära avbildningar på ändligt-dimensionella linjära rum. Vidare kan kravet på en stabilare avbildning uttryckas som så att vi vill ha ett lågt värde på operatornormen, vilket framgår av olikheten (8). Därför är det ibland av intresse att uppskatta $\|R_\alpha\|$.

Felet i approximationen (7) kan vi med utnyttjande av triangelolikheten dela upp i två termer:

$$\begin{aligned} \|x^{\alpha,\delta} - x\| &\leq \|R_\alpha y^\delta - R_\alpha y\| + \|R_\alpha y - x\| \\ &\leq \|R_\alpha\|\|y^\delta - y\| + \|R_\alpha K x - x\| \\ &\leq \delta\|R_\alpha\| + \|R_\alpha K x - x\| \end{aligned} \tag{8}$$

Detta är vår grundläggande feluppskattning. Observera att

$$\|R_\alpha K x - x\| = \|(R_\alpha - K^{-1})y\|$$

utgör approximationsfelet för det exakta högerledet y.

Felet i den regulariserade lösningen jämfört med den exakta kan följaktligen delas upp i två komponenter. Det typiska är att $\|R_\alpha\|$ växer då α går mot

noll och krymper då α går mot oändligheten. Den andra termen utgör approximationsfelet för det exakta högerledet och enligt definitionen av en regulariseringsstrategi går denna term mot noll då α går mot noll. I det typiska fallet har vi alltså en situation som illustreras av graferna i figur 4 och 5 i avsnitt 3.1.1 .

Det vi ser är att problemet att välja α kvalitativt kan beskrivas som att finna jämvikten mellan stabilitet i lösningen och närhet till den exakta lösningen i approximationen. Ett större värde på α kan resultera i en stabilare lösning, men en mindre trogen approximation. Omvänt kan ett mindre värde på α resultera i en närmare approximation, men en mindre stabil regulariserad lösning.

En särskild strategi för att välja $\alpha = \alpha(\delta, y^\delta)$, alltså beroende på δ och y^δ, kommer vi också att kalla för en *regulariseringsstrategi*, underförstått att till detta val av α korresponderar ett visst R_α. I de regulariseringsstrategier vi studerar här beror valet av α enbart på δ, så att alltså $\alpha = \alpha(\delta)$.

Definition 2.1.2. *En regulariseringsstrategi $\alpha = \alpha(\delta, y^\delta)$ säges vara* tillåten *om det för varje fixt $x \in X$ och $y \in Y$ gäller att*

$$\sup\{\alpha(\delta, y^\delta) : \|y - y^\delta\| \leq \delta,\ y^\delta \in Y\} \to 0 \text{ och}$$
$$\sup\{\|x^{\alpha(\delta,y^\delta),\delta} - x\| : \|y - y^\delta\| \leq \delta,\ y^\delta \in Y\} \to 0, \text{ då } \delta \to 0.$$

En vanlig och mycket användbar metod för att konstruera klasser av tillåtna regulariseringsstrategier är att, som man säger, *filtrera* singulärvärdessystem. Ett uttryck för lösningen x till det inversa problemet $Kx = y$ ges enligt sats 1.2.6 av

$$x = \sum_{j \in J} \frac{1}{\mu_j} \langle y, y_j \rangle x_j. \tag{9}$$

Filtreringen består nu i att i denna summa så att säga dämpa faktorerna på formen $\frac{1}{\mu_j}$ med hjälp av en funktion $q(\alpha, \mu)$.

Definition 2.1.3. *Låt $(\mu_j, x_j, y_j), j \in J$ vara ett singulärvärdessystem för K och låt $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ (eller åtminstone definierad på $\mathbb{R}^+ \times [0, \|K\|]$). För $\alpha > 0$ och $y \in Y$ sätt*

$$R_\alpha y := \sum_{j \in J} \frac{q(\alpha, \mu_j)}{\mu_j} \langle y, y_j \rangle x_j. \tag{10}$$

Om familjen $\{R_\alpha\}, \alpha > 0$, av avbildningar från Y till X utgör en regulariseringsstrategi för K säges q vara ett regulariserande filter *för K.*

Sats 2.1.1. *Låt $(\mu_j, x_j, y_j), j \in J$ vara ett singulärvärdessystem för K och låt $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ (eller åtminstone definierad på $\mathbb{R}^+ \times (0, \|K\|]$) vara en funktion med följande egenskaper:*

(1) $|q(\alpha,\mu)| \leq 1$, *för* $\alpha > 0$, $0 \leq \mu \leq \|K\|$.

(2) För alla $\alpha > 0$ *existerar* $c(\alpha)$ *sådant att* $|q(\alpha,\mu)| \leq c(\alpha)\mu$, *för* $0 \leq \mu \leq \|K\|$.

(3a) $q(\alpha,\mu) \to 1$, *då* $\alpha \to 0$, *för fixt* $0 \leq \mu \leq \|K\|$. *(Det råder alltså punktvis konvergens.)*

Då är q *ett regulariserande filter för* K *och* $\|R_\alpha\| \leq c(\alpha)$, *där regulariseringsstrategin* $\{R_\alpha\}_{\alpha>0}$, *ges av ekvation (10). Ett val* $\alpha = \alpha(\delta)$ *är tillåtet om* $\alpha(\delta) \to 0$ *och* $\delta c(\alpha(\delta)) \to 0$, *då* $\delta \to 0$.

Anmärkning. Observera att det i (3a) understrykes att konvergensen sker punktvis. Att konvergensen ej kan ske likformigt följer av villkoret (2), eftersom detta implicerar att

$$q(\alpha,\mu) \to 0, \text{ då } \mu \to 0,$$

för fixt $\alpha > 0$.

Bevis. Vi vet att $(x_j)_{j\in J}$ och $(y_j)_{j\in J}$ är ON-system i X respektive Y. Alltså kan (y_j) utvidgas till en ON-bas $(y_i)_{i\in I}$ för Y. Vi ser också att $(x_j)_{j\in J}$ är en ON-bas för X, emedan $N(K) = \{0\}$. Det följer att

$$\|y\|^2 = \sum_{i\in I} \langle y, y_i\rangle^2 \geq \sum_{j\in J} \langle y, y_j\rangle^2.$$

Från (2) och definitionen av R_α har vi att

$$\begin{aligned}\|R_\alpha y\|^2 &= \sum_{j\in J} q(\alpha,\mu_j)^2 \frac{1}{\mu_j^2} |\langle y, y_j\rangle|^2 \\ &\leq c(\alpha)^2 \sum_{j\in J} \langle y, y_j\rangle^2 \leq c(\alpha)^2 \|y\|^2.\end{aligned}$$

Alltså är $\|R_\alpha\| \leq c(\alpha)$.

Vidare är

$$R_\alpha K x = \sum_{j\in J} \frac{q(\alpha,\mu_j)}{\mu_j} \langle Kx, y_j\rangle x_j \text{ och } x = \sum_{j\in J} \langle x, x_j\rangle x_j.$$

Dessutom är $\langle Kx, y_j\rangle = \langle x, K^* y_j\rangle = \mu_j \langle x, x_j\rangle$. Det följer att

$$\begin{aligned}\|R_\alpha K x - x\|^2 &= \sum_{j\in J} \Big(\frac{q(\alpha,\mu_j)}{\mu_j} \mu_j \langle x, x_j\rangle - \langle x, x_j\rangle \Big)^2 \\ &= \sum_{j\in J} (q(\alpha,\mu_j) - 1)^2 \langle x, x_j\rangle^2. \qquad (11)\end{aligned}$$

Låt nu $x \in X$ vara fixt och låt $\epsilon > 0$. Enligt (3a) existerar $\alpha_0 > 0$ sådant att

$$(q(\alpha, \mu_j) - 1)^2 < \frac{\epsilon^2}{\|x\|^2}, \text{ för } j \in J \text{ och } 0 < \alpha \leq \alpha_0.$$

Vi drar härav slutsatsen att

$$\begin{aligned} \|R_\alpha K x - x\|^2 &= \sum_{j \in J} (q(\alpha, \mu_j) - 1)^2 \langle x, x_j \rangle^2 \\ &< \frac{\epsilon^2}{\|x\|^2} \sum_{j \in J} \langle x, x_j \rangle^2 = \frac{\epsilon^2}{\|x\|^2} \|x\|^2 = \epsilon^2, \end{aligned}$$

för alla α med $0 < \alpha \leq \alpha_0$. Alltså har vi visat att $R_\alpha K x \to x$, då $\alpha \to 0$, för alla $x \in X$.

Till det sista påståendet i satsen. Med den grundläggande feluppskattningen (8) får vi att

$$\begin{aligned} ||x^{\alpha,\delta} - x|| &\leq \delta ||R_\alpha|| + ||R_\alpha K x - x|| \\ &\leq \delta c(\alpha(\delta)) + ||R_\alpha K x - x|| \to 0, \end{aligned}$$

om $\delta c(\alpha(\delta)) \to 0$, då $\delta \to 0$. Det följer att detta villkor på valet av α ger en tillåten regulariseringsstrategi. □

Sats 2.1.2. *Låt villkoren (1) och (2) ur föregående sats gälla.*

(i) Låt (3a) ersättas av det starkare villkoret

(3b) Det existerar $c_1 > 0$ med

$$|q(\alpha, \mu) - 1| \leq c_1 \frac{\sqrt{\alpha}}{\mu}.$$

Då gäller att om $x \in K^*(Y)$, så är

$$\|R_\alpha K x - x\| \leq c_1 \sqrt{\alpha} \|z\|, \tag{12}$$

där $x = K^* z$, för något $z \in Y$.

(ii) Låt (3a) ersättas av det starkare villkoret

(3c) Det existerar $c_2 > 0$ med

$$|q(\alpha, \mu) - 1| \leq c_2 \frac{\alpha}{\mu}.$$

Då gäller att om $x \in K^* K(X)$, så är

$$\|R_\alpha K x - x\| \leq c_2 \alpha \|z\|, \tag{13}$$

där $x = K^* K z$, för något $z \in X$.

Bevis. Fall (i): Låt $x = K^*z$, där $z \in Y$. Då är $\langle x, x_j\rangle = \mu_j\langle z, y_j\rangle$ och ekvation (11) ger att

$$\begin{aligned}\|R_\alpha Kx - x\|^2 &= \sum_{j\in J}(q(\alpha,\mu_j)-1)^2\mu_j^2\langle z, y_j\rangle^2 \\ &\le c_1^2\alpha\sum_{j\in J}\langle z, y_j\rangle^2 \le c_1^2\alpha\|z\|^2,\end{aligned}$$

det vill säga $\|R_\alpha Kx - x\| \le c_1\sqrt{\alpha}\|z\|$.

Fall (ii): Visas på analogt sätt. Vi har att $\langle x, x_j\rangle = \langle K^*Kz, x_j\rangle = \langle z, K^*Kx_j\rangle$, eftersom K^*K är symmetrisk, och $K^*Kx_j = K^*(\mu_j y_j) = \mu_j(K^*y_j) = \mu_j^2 x_j$. Alltså är $\langle x, x_j\rangle = \mu_j^2\langle z, x_j\rangle$. Ekvation (11) ger att

$$\begin{aligned}\|R_\alpha Kx - x\|^2 &= \sum_{j\in J}(q(\alpha,\mu_j)-1)^2\langle x, x_j\rangle^2 \\ &\le \sum_{j\in J}\left(c_2\frac{\alpha}{\mu_j^2}\right)^2(\mu_j^2\langle z, x_j\rangle)^2 = c_2^2\alpha^2\sum_{j\in J}\langle z, x_j\rangle^2 \\ &\le c_2^2\alpha^2\|z\|^2,\end{aligned}$$

det vill säga $\|R_\alpha Kx - x\| \le c_2\alpha\|z\|$. □

Förutsätter vi att K är injektiv kan villkoren på x i ovanstående sats förenklas:

Sats 2.1.3. *Om den linjära avbildningen K är injektiv, så är*

$$K^*(Y) = K^*K(X) = X.$$

Bevis. Det räcker förstås med att visa att $K^*K(X) = X$. Eftersom K^*K är en linjär operator på det euklidiska rummet X är vi klara om vi kan visa att K^*K är injektiv. Surjektiviteten följer ju då av injektiviteten. Antag alltså att

$$K^*Kx_1 = K^*Kx_2,$$

där $x_1,\ x_2 \in X$. Detta innebär att för alla $z \in X$ är

$$\langle Kx_1, Kz\rangle = \langle Kx_2, Kz\rangle.$$

Sätter vi $z = x_1 - x_2$ får vi att

$$\begin{aligned}&\langle K(x_1-x_2), K(x_1-x_2)\rangle = 0 \\ \Longrightarrow\quad &K(x_1-x_2) = 0 \\ \Longrightarrow\quad &x_1 = x_2.\end{aligned}$$

□

Det finns många exempel på regulariserande filter som uppfyller egenskaperna (1), (2) och (3a-c). Vi nöjer oss med att ge två för oss särskilt relevanta exempel:

Sats 2.1.4. *Följande två funktioner uppfyller villkoren (1), (2) och (3a-c) ur föregående satser, och är alltså regulariserande filter:*

(i)

$$q(\alpha, \mu) = \frac{\mu^2}{\alpha + \mu^2}.$$

Denna uppfyller villkor (2) med $c(\alpha) = 1/(2\sqrt{\alpha})$. *Villkor (3b) och (3c) gäller med* $c_1 = 1/2$ *och* $c_2 = 1$ *respektive.*

(ii)

$$q(\alpha, \mu) = 1 - (1 - a\mu^2)^{1/\alpha}, \text{ där } 0 < a < 1/\|K\|^2.$$

I detta fall gäller villkoren med $c(\alpha) = \sqrt{a/\alpha}$, $c_1 = 1/\sqrt{2a}$ *och* $c_2 = 1/a$.

Bevis. Att båda funktionerna uppfyller egenskaperna (1) och (3a) är klart.

Fall (i): Då $0 \leq (\sqrt{\alpha} - \mu)^2$ har vi att

$$\frac{\mu}{\alpha + \mu^2} \leq \frac{1}{2\sqrt{\alpha}},$$

varur villkor (2) följer med $c(\alpha) = 1/(2\sqrt{\alpha})$. Vidare är

$$1 - q(\alpha, \mu) = \frac{\alpha}{\alpha + \mu^2},$$

och ur ovanstående olikhet får vi att

$$\alpha \frac{\mu}{\alpha + \mu^2} \leq \frac{1}{2\sqrt{\alpha}} = 1/2\sqrt{\alpha}.$$

Alltså är

$$\frac{\alpha}{\alpha + \mu^2} \leq 1/2 \frac{\sqrt{\alpha}}{\mu},$$

så villkor (3b) gäller med $c_1 = 1/2$. Villkor (3c) följer, med $c_2 = 1$, av att

$$\frac{\alpha}{\alpha + \mu^2} \leq \frac{\alpha}{\mu^2}.$$

Fall (ii): MacLaurinutveckling av funktionen $f(t) = (1 - t)^{1/\alpha}$ ger

$$f(t) = 1 - \frac{t}{\alpha} + R_2(t)t^2,$$

där den sista termen är Lagranges restterm. För $0 \leq t \leq 1$ är resttermen icke-negativ, varav följer att

$$(1 - a\mu^2)^{1/\alpha} \geq 1 - \frac{a\mu^2}{\alpha},$$

för $0 \leq a\mu^2 \leq 1$. Alltså är

$$1 - (1 - a\mu^2)^{1/\alpha} \leq 1 - (1 - \frac{a\mu^2}{\alpha}) = \frac{a\mu^2}{\alpha}.$$

Vidare är

$$|q(\alpha, \mu)| \leq \sqrt{|q(\alpha, \mu)|} \leq \sqrt{\frac{a}{\alpha}}\mu.$$

Villkor (2) följer med $c(\alpha) = \sqrt{a/\alpha}$.

Sätt nu

$$f(t) = t(1 - t)^\beta, \quad \beta > 0.$$

Optimering av $f(t)$ på intervallet $[0, 1]$ ger att funktionens maximum antas i punkten $t = 1/\beta$, vilket i sin tur ger att

$$f(t) \leq f(\frac{1}{\beta}) = \frac{1}{\beta}(1 - \frac{1}{\beta})^\beta \leq \frac{1}{\beta},$$

på det relevanta intervallet. Av detta följer att, för $0 \leq t \leq 1/a$,

$$\begin{aligned} at(1 - at)^\beta \leq \frac{1}{\beta} \implies \\ t(1 - at)^\beta \leq \frac{1}{a\beta}, \end{aligned}$$

vilket ger att

$$\mu^2(1 - a\mu^2)^\beta \leq \frac{1}{a\beta}, \tag{14}$$

för $0 \leq \mu \leq 1/\sqrt{a}$. Med $\beta = 1/\alpha$ och $c_2 = 1/a$ följer (3c) av denna olikhet.

Från (14) får vi att

$$\begin{aligned} \mu^2(1 - a\mu^2)^{2\beta} \leq \frac{1}{2a\beta} \implies \\ \mu(1 - a\mu^2)^\beta \leq \frac{1}{\sqrt{2a\beta}}, \end{aligned}$$

för $0 \leq \mu \leq 1/\sqrt{a}$, varav analogt (3b) följer med $c_1 = 1/\sqrt{2a}$. □

I de kommande två avsnitten ges ekvivalenta framställningar eller formuleringar av till dessa filter motsvarande regulariseringsstrategier, och för bevis av relevanta satser utnyttjas resultaten i detta avsnitt. Skillnaden består i att dessa framställningar i sig inte formuleras i termer av filtrerade singulärvärdessystem, varför vi alltså inte är beroende av kännedom om vare sig singulärvärdena eller något singulärvärdessystem. Dessa enklare formuleringar är de som senare används i de numeriska experimenten.

2.2 Tikhonovregularisering

En vanlig metod för att hantera linjära ekvationssystem $Kx = y$ utan lösning är att försöka finna en "anpassad" lösning ("the best fit") till problemet, genom att minimera uttrycket $\|Kx - y\|$, eller ekvivalent $\|Kx - y\|^2$, med avseende på $x \in X$. För euklidiska rum är detta problem alltid lösbart:

Sats 2.2.1. *Minimeringsproblemet*

$$\|Kx - y\| \tag{15}$$

är lösbart med avseende på $x \in X$ för varje fixt $y \in Y$. Det existerar med andra ord ett $x_0 \in X$ sådant att $\|Kx_0 - y\| \leq \|Kx - y\|$, för alla $x \in X$. Varje lösning x_0 till minimeringsproblemet (15) är en lösning till ekvationen

$$K^*Kx = K^*y \tag{16}$$

och omvänt. Om K vidare är injektiv, så är detta minimum entydigt bestämt.

Bevis. Låt oss börja med att visa att minimeringsproblemet är lösbart.

Fall 1. K är injektiv. Sätt

$$c = \min_{\|x\|=1} \|Kx\|.$$

Detta minimivärde existerar och är större än noll eftersom $\|Kx\|$ är kontinuerlig på den kompakta sfären $\{\|x\| = 1 : x \in X\}$ och K vidare är en injektiv linjär avbildning. Vi har att

$$\|Kx\| = \|x\|\|K(x/\|x\|)\| \geq \|x\|c \to \infty, \text{ då } \|x\| \to \infty.$$

Av denna observation följer att

$$\|Kx - y\| \to \infty, \text{ då } \|x\| \to \infty.$$

Alltså kan vi välja ett slutet klot B kring 0 sådant att $\|Kx - y\| \geq C$ för alla x utanför B, där $C \in \mathbb{R}$. Om ett minimum existerar ligger det alltså i B. Existensen av ett minimum följer här av att B är en kompakt mängd i X och att $\|Kx - y\|$ är kontinuerlig. Vi kan också konstatera att detta minimum är entydigt bestämt, eftersom funktionen $\|Kx - y\|^2$ är strängt konvex.

Fall 2. K är ej injektiv. Då ersätter vi definitionsrummet X med delrummet $N(K)^\perp$. På detta delrum är K injektiv och har ett minimum, som vidare är entydigt bestämt.

Till det andra påståendet i satsen. *Binomialformeln* för linjära rum med skalärprodukt säger att

$$\|x \pm y\|^2 = \|x\|^2 + \|y\|^2 \pm 2\langle x, y\rangle,$$

för alla $x, y \in X$. Med användande av denna får vi att

$$\begin{aligned}\|Kx - y\|^2 &= \|(Kx_0 - y) + K(x - x_0)\|^2 \\ &= \|Kx_0 - y\|^2 + \|K(x - x_0)\|^2 + 2\langle Kx_0 - y, K(x - x_0)\rangle.\end{aligned}$$

för alla $x, x_0 \in X$. Alltså är

$$\begin{aligned}\|Kx - y\|^2 - \|Kx_0 - y\|^2 &= 2\langle Kx_0 - y, K(x - x_0)\rangle + \|K(x - x_0)\|^2 \\ &= 2\langle K^*(Kx_0 - y), x - x_0\rangle + \|K(x - x_0)\|^2. \quad (17)\end{aligned}$$

Antag att x_0 satisfierar ekvation (16), så att $K^*(Kx_0 - y) = 0$. Då är

$$\|Kx - y\|^2 - \|Kx_0 - y\|^2 \geq 0,$$

för alla $x \in X$, och alltså är x_0 en lösning till minimeringsproblemet.

Antag, omvänt, att x_0 är en lösning till minimeringsproblemet. Låt $z \in X$ vara godtyckligt. Betrakta den räta halvlinjen $x = x_0 + tz$, $t > 0$. Likheten (17) ger att

$$0 \leq 2t\langle K^*(Kx_0 - y), z\rangle + t^2\|Kz\|^2.$$

Delar vi högerledet med t och låter $t \to 0$ får vi att

$$0 \leq \langle K^*(Kx_0 - y), z\rangle,$$

för alla $z \in X$. Detta är endast möjligt om $K^*(Kx_0 - y) = 0$. □

Låt oss nu återigen betrakta den perturberade ekvationen (4),

$$Kx^\delta = y^\delta.$$

Oavsett om detta problem är väl ställt eller ej kunde vi försöka finna en anpassad lösning genom att lösa ovanstående minimeringsproblem. Det är förvisso klart att $\|Kx - y\|$ inte ligger långt från $\|Kx^\delta - y^\delta\|$, eftersom

$$\big|\|Kx - y\| - \|Kx - y^\delta\|\big| \leq \|(Kx - y) - (Kx - y^\delta)\| = \|y - y^\delta\| \leq \delta,$$

för alla $x \in X$. Nackdelen med detta förfarande är att den resulterande lösningen x^δ inte nödvändigtvis ligger särskilt nära x, så att $\|x - x^\delta\|$ kan vara betydligt större än δ. Låt exempelvis K vara avbildningen

$$Kx = ax, \ a > 0, \ x \in \mathbb{R}.$$

Detta ger

$$\|x - x^\delta\| = \|\frac{y}{a} - \frac{y^\delta}{a}\| = \frac{1}{a}\|y - y^\delta\| \leq \frac{\delta}{a},$$

vilket innebär ett alltför stort spelrum för felet i lösningen i det fall a är mycket litet.

En väg ut ur detta dilemma är att så att säga bromsa, hämma, dämpa x^δ genom att i minimeringsproblemet lägga ytterligare en term till funktionalen $\|Kx - y\|^2$, som enbart beror på x.

Tikhonovs minimeringsproblem: Givet den linjära avbildningen $K : X \longrightarrow Y$ och ett $y \in Y$, består problemet i att finna ett $x^\alpha \in X$ som minimerar den så kallade *Tikhonovfunktionalen*

$$J_\alpha(x) = J_\alpha(x;\ y) := \|Kx - y\|^2 + \alpha\|x\|^2, \quad x \in X,\ \alpha > 0.$$

För godtyckligt $y \in Y$ definierar vi nu

$$R_\alpha(y) := \text{argmin}_{x\in X}\, J_\alpha(x).$$

där argmin ger det argument till J_α för vilket funktionen antar sitt minimivärde (eller mer generellt, mängden av sådana argument). Följande sats säger oss att denna definition faktiskt ger en väldefinierad avbildning från Y till X för varje $\alpha > 0$.

Sats 2.2.2. *Tikhonovfunktionalen J_α har för varje $\alpha > 0$ ett entydigt bestämt minimum $x^\alpha \in X$, vilket också är den entydiga lösningen till den så kallade* normalekvationen

$$\alpha x^\alpha + K^*Kx^\alpha = K^*y. \tag{18}$$

Lösningen x^α till normalekvationen kan följaktligen skrivas på formen $x^\alpha = R_\alpha y$, med

$$R_\alpha = (\alpha I + K^*K)^{-1}K^*, \tag{19}$$

för varje $\alpha > 0$.

Bevis. Låt $I = \inf\{J_\alpha(x) : x \in X\}$ och låt $(x_n)_1^\infty$ vara en minimerande följd i X. Låt med andra ord $J_\alpha(x_n) \to I$, då $n \to \infty$. Vi ska visa att (x_n) är en Cauchyföljd i X. Binomialformeln kan skrivas om på formen

$$\|x\|^2 + \|y\|^2 = \frac{1}{2}\|x + y\|^2 + \frac{1}{2}\|x - y\|^2.$$

Med utnyttjande av denna omskrivning får vi att

$$\begin{aligned} J_\alpha(x_n) + J_\alpha(x_m) &= 2J_\alpha(\frac{1}{2}(x_n + x_m)) + \frac{1}{2}\|K(x_n - x_m)\|^2 + \frac{1}{2}\alpha\|x_n - x_m\|^2 \\ &\geq 2I + \frac{1}{2}\alpha\|x_n - x_m\|^2. \end{aligned}$$

Vänsterledet i olikheten konvergerar mot $2I$ då $n,\ m \to \infty$, varav följer att $\|x_n - x_m\| \to 0$, då då $n,\ m \to \infty$. Detta visar att (x_n) är en Cauchyföljd och alltså konvergent i X. Sätt

$$x^\alpha = \lim_{n\to\infty} x_n.$$

Eftersom J_α är kontinuerlig drar vi slutsatsen att $J_\alpha(x_n) \to J_\alpha(x^\alpha)$, då $n \to \infty$, och alltså att $J_\alpha(x^\alpha) = I$. Detta visar att J_α har åtminstone en minimipunkt.

Ekvivalensen mellan normalekvationen (18) och Tikhonovs minimeringsproblem visas på samma sätt som i beviset för sats 2.2.1, där vi kan utnyttja följande likhet:

$$\begin{aligned} J_\alpha(x) - J_\alpha(x^\alpha) &= 2\langle Kx^\alpha - y, K(x - x^\alpha)\rangle + 2\alpha\langle x^\alpha, x - x^\alpha\rangle \\ &\quad + \|K(x - x^\alpha)\|^2 + \alpha\|x - x^\alpha\|^2 \\ &= 2\langle K^*(Kx^\alpha - y) + \alpha x^\alpha, x - x^\alpha\rangle \\ &\quad + \|K(x - x^\alpha)\|^2 + \alpha\|x - x^\alpha\|^2, \end{aligned}$$

för alla $x \in X$.

Vi ska slutligen visa entydigheten av x^α. Detta gör vi genom att visa att den linjära operatorn $\alpha I + K^*K : X \longrightarrow X$ är injektiv för varje $\alpha > 0$. Antag att $\alpha x + K^*Kx = 0$. Skalärmultiplikation med x ger att

$$\langle \alpha x + K^*Kx, x\rangle = \alpha\langle x, x\rangle + \langle Kx, Kx\rangle = 0,$$

varav följer att $x = 0$. Alltså är $\alpha I + K^*K$ injektiv. □

Familjen $\{R_\alpha\}_{\alpha>0}$ av avbildningar är alltså *Tikhonovs regulariseringsstrategi*. Användande av resultaten i föregående avsnitt ger

Sats 2.2.3. *Låt $K : X \longrightarrow Y$ vara en linjär avbildning. Då gäller att:*

(i) Familjen $\{R_\alpha : Y \longrightarrow X\}_{\alpha>0}$ av avbildningar utgör en regulariseringsstrategi för K, med $\|R_\alpha\| \leq 1/(2\sqrt{\alpha})$. Varje val $\alpha = \alpha(\delta)$ är tillåtet om $\alpha(\delta) \to 0$ och $\delta^2/\alpha(\delta) \to 0$, då $\delta \to 0$.

(ii) Låt $x = K^*z$, $z \in Y$, med $\|z\| \leq E$ för något $E > 0$. Med valet $\alpha(\delta) = c_0\delta/E$, för något $c_0 > 0$, får vi feluppskattningen

$$\|x^{\alpha(\delta),\delta} - x\| \leq \frac{1}{2}\left(\frac{1}{\sqrt{c_0}} + \sqrt{c_0}\right)\sqrt{\delta E}. \tag{20}$$

(iii) Låt $x = K^*Kz$, $z \in X,$, med $\|z\| \leq E$ för något $E > 0$. Med valet $\alpha(\delta) = c_0(\delta/E)^{2/3}$, för något $c_0 > 0$, får vi feluppskattningen

$$\|x^{\alpha(\delta),\delta} - x\| \leq \left(\frac{1}{2\sqrt{c_0}} + c_0\right)E^{1/3}\delta^{2/3}. \tag{21}$$

Bevis. (i): Låt $(\mu_j, x_j, y_j)_{j\in J}$ vara ett singulärvärdessystem för K. Vi ska visa att

$$x^\alpha = R_\alpha y = \sum_{j\in J} \frac{\mu_j}{\alpha + \mu_j^2}\langle y, y_j\rangle x_j. \tag{22}$$

Detta gör vi genom att sätta in högerledet i denna ekvation i vänsterledet till ekvation (18). Till att börja med har vi att

$$\begin{aligned} K^*K\Big(\sum_{j\in J}\frac{\mu_j}{\alpha+\mu_j^2}\langle y,y_j\rangle x_j\Big) &= K^*\Big(\sum_{j\in J}\frac{\mu_j}{\alpha+\mu_j^2}\langle y,y_j\rangle Kx_j\Big)\\ &= \sum_{j\in J}\frac{\mu_j^2}{\alpha+\mu_j^2}\langle y,y_j\rangle K^*y_j\\ &= \sum_{j\in J}\frac{\mu_j^3}{\alpha+\mu_j^2}\langle y,y_j\rangle x_j. \end{aligned}$$

Alltså är enligt singulärvärdessatsen

$$\begin{aligned} &\alpha\sum_{j\in J}\frac{\mu_j}{\alpha+\mu_j^2}\langle y,y_j\rangle x_j + K^*K\Big(\sum_{j\in J}\frac{\mu_j}{\alpha+\mu_j^2}\langle y,y_j\rangle x_j\Big)\\ &= \sum_{j\in J}\frac{\alpha\mu_j}{\alpha+\mu_j^2}\langle y,y_j\rangle x_j + \sum_{j\in J}\frac{\mu_j^3}{\alpha+\mu_j^2}\langle y,y_j\rangle x_j\\ &= \sum_{j\in J}\mu_j\langle y,y_j\rangle x_j = K^*y. \end{aligned}$$

Med

$$q(\alpha,\mu)=\frac{\mu^2}{\alpha+\mu^2},\quad \alpha,\ \mu>0,$$

ser vi följaktligen att R_α har representationen

$$R_\alpha y=\sum_{j\in J}\frac{\mu_j}{\alpha+\mu_j^2}\langle y,y_j\rangle x_j=\sum_{j\in J}\frac{q(\alpha,\mu_j)}{\mu_j}\langle y,y_j\rangle x_j, \tag{23}$$

för alla $y\in Y$. Funktionen q är det regulariserande filtret i sats 2.1.4 del (i). Av den satsen och sats 2.1.1 följer att $\{R_\alpha\}_{\alpha>0}$ är en regulariseringsstrategi med $\|R_\alpha\|\le 1/(2\sqrt{\alpha})$ och att valet $\alpha=\alpha(\delta)$ är tillåtet om $\alpha(\delta)\to 0$ och $\delta c(\alpha(\delta))=\delta/(2\sqrt{\alpha(\delta)})\to 0$ då $\delta\to 0$. Det senare är förstås ekvivalent med att $\delta^2/\alpha(\delta)\to 0$ då $\delta\to 0$.

(ii): Den allmänna feluppskattningen (8), sats 2.1.2 del (i) och sats 2.1.4 del (i) ger tillsammans att

$$\begin{aligned} \|x^{\alpha(\delta),\delta}-x\| &\le \delta\|R_\alpha\|+\|R_\alpha Kx-x\| \le \frac{\delta}{2\sqrt{\alpha(\delta)}}+\frac{1}{2}\sqrt{\alpha(\delta)}\|z\|\\ &\le \frac{\delta}{2\sqrt{\alpha(\delta)}}+\frac{1}{2}\sqrt{\alpha(\delta)}E=\frac{1}{2}\Big(\frac{1}{\sqrt{c_0}}+\sqrt{c_0}\Big)\sqrt{\delta E}. \end{aligned}$$

(iii): Med sats 2.1.2 del (ii) istället för del (i) får vi i analogi med ovanstående fall att

$$\|x^{\alpha(\delta),\delta} - x\| \leq \delta\|R_\alpha\| + \|R_\alpha Kx - x\| \leq \frac{\delta}{2\sqrt{\alpha(\delta)}} + \alpha(\delta)\|z\|$$
$$\leq \frac{\delta}{2\sqrt{\alpha(\delta)}} + \alpha(\delta)E = \Big(\frac{1}{2\sqrt{c_0}} + c_0\Big)E^{1/3}\delta^{2/3}. \qquad \square$$

2.3 Generaliserad Tikhonovregularisering

En möjlig generalisering av Tikhonovfunktionalen J_α är funktionalen

$$G_\alpha(x) = G_\alpha(x;\ y) := \frac{1}{2}(\|Kx - y\|^2 + \alpha\langle Lx, x\rangle), \tag{24}$$

där $x \in X$, $y \in Y$ är fixt och L är en linjär operator på X. Vi förutsätter här att L är symmetrisk och strängt positiv. Låt oss definiera det sistnämnda begreppet tillsammans med några ytterligare:

Definition 2.3.1. *Låt A vara en linjär operator på X. Sätt*

$$\lambda_{\min}(A) := \inf_{\|x\|=1} \langle Ax, x\rangle$$

Vi säger att A är positivt semidefinit *om $\lambda_{min}(A) \geq 0$ och* strängt positiv *om sträng olikhet råder.*

Vi säger också att en linjär operator A är *positivt definit* om

$$\langle Ax, x\rangle > 0,$$

för alla $x \neq 0$, $x \in X$. I det ändligt-dimensionella fallet är detta dock ekvivalent med att A är strängt positiv. Vidare är förstås

$$\inf_{\|x\|=1} \langle Ax, x\rangle = \min_{\|x\|=1} \langle Ax, x\rangle.$$

Sats 2.3.1. *Låt B vara en linjär operator på X. Uttrycket*

$$\|Kx - y\|^2 + \alpha\|Bx\|^2 \tag{25}$$

kan då skrivas om på formen (24), för någon positivt semidefinit symmetrisk operator L. Omvänt, givet den positivt semidefinita symmetriska operatorn L existerar det en linjär operator B på X sådan att (24) kan skrivas om på formen (25).

Anmärkning. Villkoret att B är en linjär operator kan försvagas lite. Det går också bra med en linjär avbildning $B : X \longrightarrow Z$, där Z är ett euklidiskt rum, för vilken vi liksom i beviset sätter $L = B^*B$.

Bevis. För den första implikationen sätt $L = B^*B$. Vi har då att

$$\|Bx\|^2 = \langle Bx, Bx\rangle = \langle B^*Bx, x\rangle = \langle Lx, x\rangle.$$

För den andra implikationen, låt $\lambda_1, \ldots, \lambda_n$ vara (den ordnade) följden av egenvärden till L räknade med multiplicitet. Eftersom L är symmetrisk är matrisen $[L]$ diagonaliserbar (i godtycklig ON-bas för X) , med

$$[L] = T^tDT$$

för någon $n \times n$-matris T, där $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Eftersom L är positivt semidefinit är alla egenvären till L icke-negativa. Alltså kan vi bilda en linjär operator B med

$$[B] = \sqrt{D}T$$

(i den givna ON-basen för X). Vi får att

$$[B^*B] = [B]^t[B] = T^t(\sqrt{D})^2T = [L].$$

Alltså är $L = B^*B$. □

Nu låter vi för ett godtyckligt $y \in Y$

$$R_\alpha(y) := \text{argmin}_{x\in X}\, G_\alpha(x)$$

(där argmin definieras såsom i föregående avsnitt). Vår målsättning är att visa att denna definition ger en väldefinierad avbildning från Y till X och att familjen $\{R_\alpha\}_{\alpha>0}$ utgör en linjär regulariseringsstrategi för K. För detta utnyttjar vi en del elementär konvexitetsteori.

Definition 2.3.2. *Låt $f : X \longrightarrow \mathbb{R}$. Vi säger att f är* konvex *om*

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2),$$

för alla $x_1,\ x_2 \in X$ och $\theta \in [0,1]$. Funktionen f säges vara strängt konvex *om olikheten är sträng närhelst $x_1 \neq x_2$ och $\theta \in (0,1)$.*

Sats 2.3.2. *Om $f : X \longrightarrow \mathbb{R}$ är konvex, så är varje lokal minimipunkt till f global. Om f är strängt konvex och har ett minimivärde, så är minimipunkten entydigt bestämd.*

Bevis. Låt x_0 vara en lokal minimipunkt till f. Vi har då att $f(z) \geq f(x_0)$ för alla $z \in U$, där U är någon omgivning till x_0. För ett godtyckligt $x \in X$ gäller att punkten $z = \theta x_0 + (1-\theta)x$ ligger i U, för något $0 < \theta < 1$ tillräckligt nära 1. Alltså är

$$\theta f(x_0) + (1-\theta)f(x) \geq f(\theta x_0 + (1-\theta)x) = f(z) \geq f(x_0),$$

vilket ger att $f(x) \geq f(x_0)$. Därav följer globaliteten.

Antag nu att f är strängt konvex och att x_0 och x_1 är minimipunkter. Vi ska visa att $x_0 = x_1$. Antag det motsatta, att x_0 och x_1 är distinkta punkter i X. Låt $0 < \theta < 1$. Då är

$$f(\theta x_0 + (1-\theta)x_1) < \theta f(x_0) + (1-\theta)f(x_1) = f(x_0),$$

en motsägelse. □

Den naturliga normen $||\cdot||$ på ett linjärt rum med skalärprodukt är konvex och dess kvadrat, alltså funktionen $||\cdot||^2$, är strängt konvex. Detta är enkla konsekvenser av definitionen av normen och räknelagarna för skalärprodukten. Låt $K : X \longrightarrow Y$ vara en injektiv linjär avbildning och låt $y \in Y$ vara fixt. Då är funktionen $\|Kx - y\|^2$ strängt konvex. Låt nämligen x_1 och x_2 vara distinkta pukter i X och låt $0 < \theta < 1$. Vi har att

$$\begin{aligned}\|K(\theta x_1 + (1-\theta)x_2) - y\|^2 &= \|\theta(Kx_1 - y) + (1-\theta)(Kx_2 - y)\|^2 \\ &< \theta\|Kx_1 - y\|^2 + (1-\theta)\|Kx_2 - y\|^2,\end{aligned}$$

eftersom $Kx_1 - y \neq Kx_2 - y$.

Sats 2.3.3. *Låt G_α vara skriven på formen (25). Då är G_α strängt konvex om $N(K) \cap N(B) = \{0\}$.*

Bevis. Vi påminner oss om att funktionen $||\cdot||^2$ är strängt konvex. Låt nu $x_1 \neq x_2$ vara punkter i X och låt $0 < \theta < 1$. Vi får då att

$$\begin{aligned}G_\alpha(\theta x_1 + (1-\theta)x_2) &= ||K(\theta x_1 + (1-\theta)x_2) - y||^2 + \alpha||B(\theta x_1 + (1-\theta)x_2)||^2 \\ &= ||\theta(Kx_1 - y) + (1-\theta)(Kx_2 - y)||^2 \\ &\quad + \alpha\|\theta Bx_1 + (1-\theta)Bx_2\|^2 \\ &\leq (\theta\|Kx_1 - y\|^2 + (1-\theta)\|Kx_2 - y\|^2) \\ &\quad + \alpha(\theta\|Bx_1\|^2 + (1-\theta)\|Bx_2\|^2) \\ &= \theta G_\alpha(x_1) + (1-\theta)G_\alpha(x_2),\end{aligned}$$

med sträng olikhet om $Kx_1 - y \neq Kx_2 - y$ eller $Bx_1 \neq Bx_2$. Detta är ekvivalent med att $K(x_1 - x_2) \neq 0$ eller $B(x_1 - x_2) \neq 0$, vilket är fallet eftersom x_1 och x_2 är distinkta och vi antar att $N(K) \cap N(B) = \{0\}$. □

Om K är injektiv är förstås villkoret i satsen uppfyllt. Vi är nu redo att visa att

Sats 2.3.4. *R_α är en väldefinierad avbildning från Y till X.*

Bevis. Vi behöver ett litet hjälpresultat för detta bevis: det existerar $c_0 > 0$ sådant att

$$\langle Lx, x\rangle \geq c_0\|x\|^2, \tag{26}$$

för alla $x \in X$. Enligt antagande är L strängt positiv. Sätt $c_0 = \lambda_{\min}(L)$ och låt $x \in X$ vara godtyckligt. Vi kan skriva x på formen $x = ax_0$, för något $a \in \mathbb{R}$ och $x_0 \in X$ med $\|x_0\| = 1$. Vi har då att

$$\begin{aligned}\langle Lx, x\rangle &= \langle L(ax_0), ax_0\rangle = a^2\langle Lx_0, x_0\rangle \\ &\geq a^2 c_0 = a^2 c_0 \|x_0\|^2 = c_0\|ax_0\|^2 = c_0\|x\|^2.\end{aligned}$$

Alltså gäller (26).

Av sats 2.3.3 följer att G_α är en strängt konvex funktional från X till $\mathbb{R}$ för varje fixt $y \in Y$. Av sats 2.3.2 följer att en eventuell minimipunkt är entydigt bestämd. Nu gäller det alltså bara att visa att minimum faktiskt existerar, för att vi ska kunna dra slutsatsen att R_α är väldefinierad. (Observera att sträng konvexitet inte implicerar att ett minimum existerar.) Av olikheten (26) följer att $G_\alpha(x) \to +\infty$ då $\|x\| \to \infty$. Alltså kan vi välja ett slutet klot B kring 0 sådant att $G_\alpha(x) \geq C$ för alla x utanför B, där $C \in \mathbb{R}$. Om ett minimum existerar ligger det alltså i B. Existensen av ett minimum följer här av att B är en kompakt mängd i X och att G_α är kontinuerlig. □

Sats 2.3.5. *Avbildningarna R_α konvergerar punktvis mot identitetsoperatorn på X, d v s,*

$$R_\alpha K x \to x, \text{ då } \alpha \to 0,$$

för alla $x \in X$.

Bevis. Låt $x_0 \in X$ vara godtyckligt men fixt. Vi har att

$$G_\alpha(x;\ Kx_0) = \|Kx - Kx_0\|^2 + \alpha\langle Lx, x\rangle$$

Den andra termen går mot noll, då α går mot noll, och den första termen är oberoende av α. Eftersom $\|Kx_0 - Kx_0\|^2 = 0$ gäller alltså att

$$\lim_{\alpha\to 0} R_\alpha K x = \lim_{\alpha\to 0} \operatorname{argmin}_{x\in X} G_\alpha(x;\ Kx_0) = x_0. \qquad \square$$

Det återstår för oss att visa att R_α är kontinuerlig. Vi påminner oss till att börja med om sambandet mellan riktningsderivata och gradient (för bevis hänvisas läsaren till någon lärobok i flervariabel analys).

Sats 2.3.6. *Om $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ är differentierbar i punkten $x \in \mathbb{R}^n$ och v är en vektor i $\mathbb{R}^n$, så existerar riktningsderivatan $\delta_v f(x)$ av f i punkten x i riktningen v. Dessutom är*

$$\delta_v f(x) = \frac{d}{d\tau} f(x + \tau v)\Big|_{\tau=0} = \langle \nabla f(x), v\rangle.$$

Sats 2.3.7. *Låt $x_0 \in \mathbb{R}^n$ vara en lokal minimipunkt till funktionen $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ och låt U vara en konvex minimiomgivning till x_0 för f. Antag vidare att f är differentierbar i x_0. Då är*

$$\langle \nabla f(x_0), x - x_0\rangle \geq 0, \tag{27}$$

för alla $x \in U$.

Bevis. Låt $x \in U$ vara godtyckligt. Då U är konvex ligger även punkten $x_0 + \tau(x - x_0) = \tau x + (1-\tau)x_0$ i U, för $0 \leq \tau \leq 1$. Eftersom x_0 är ett lokalt minimum är

$$\lim_{\tau \to 0+} \frac{f(x_0 + \tau(x - x_0)) - f(x_0)}{\tau} \geq 0,$$

där vi i högerledet har riktningsderivatan av f i punkten x_0 i riktningen $(x - x_0)$. Olikheten (27) följer av sats 2.3.6. □

Dessa satser kan genom isomorfi översättas till motsvarande satser för allmänna euklidiska rum.

Sats 2.3.8. *Avbildningen R_α är kontinuerlig.*

Bevis. Till att börja med ska vi beräkna gradienten till G_α för ett fixt $y \in Y$. För varje $v \in X$ och $\tau \in \mathbb{R}$ gäller att

$$\begin{aligned}
G_\alpha(x + \tau v) &= \frac{1}{2}\|K(x + \tau v) - y\|^2 + \frac{1}{2}\alpha\langle L(x + \tau v), x + \tau v\rangle \\
&= \frac{1}{2}(\|Kx - y\|^2 + \tau^2\|Kv\|^2 + 2\tau\langle Kx - y, Kv\rangle) \\
&\quad + \frac{1}{2}\alpha(\langle Lx, x\rangle + \tau\langle Lx, v\rangle + \tau\langle Lv, x\rangle + \tau^2\langle Lv, v\rangle) \\
&= (\frac{1}{2}\|Kx - y\|^2 + \frac{1}{2}\alpha\langle Lx, x\rangle) \\
&\quad + \tau(\langle Kx - y, Kv\rangle) + \alpha\langle Lx, v\rangle) \\
&\quad + \tau^2(\frac{1}{2}\|Kv\|^2 + \frac{1}{2}\alpha\langle Lv, v\rangle).
\end{aligned}$$

(Här har vi återinfört skalfaktorn $\frac{1}{2}$ i det ursprungliga uttrycket för den generaliserade tikhonovfunktionalen för att bli av med en del 2:or som annars skulle finnas med i uttrycket för gradienten.) Derivering med avseende på τ ger att

$$\delta_v G_\alpha(x) = \langle Kx - y, Kv\rangle + \alpha\langle Lx, v\rangle = \langle K^*(Kx - y) + \alpha Lx, v\rangle.$$

Av sats 2.3.6 följer att

$$\nabla G_\alpha(x) = (K^*K + \alpha L)x - K^*y. \tag{28}$$

Låt nu $y_0 \in Y$ vara fixt och låt följden $y_n \to y_0$, $n = 1, 2, 3, \ldots$. Sätt $x_0 = R_\alpha(y_0)$ och $x_n = R_\alpha(y_n)$, $n = 1, 2, 3, \ldots$. För att förenkla beteckningarna sätter vi $G_0(x) = G_\alpha(x; y_0)$ och $G_n(x) = G_\alpha(x; y_n)$, $n = 1, 2, 3, \ldots$. Från sats 2.3.7 har vi att

$$\begin{aligned}
&\langle \nabla G_n(x_n), x_0 - x_n\rangle \geq 0, \quad \text{varför} \\
&\langle \nabla G_n(x_n), x_n - x_0\rangle \leq 0, \quad \text{och att} \\
&\langle \nabla G_0(x_0), x_n - x_0\rangle \geq 0.
\end{aligned}$$

Alltså är

$$\begin{aligned} 0 &\geq \langle \nabla G_n(x_n), x_n - x_0 \rangle - \langle \nabla G_0(x_0), x_n - x_0 \rangle \\ &= \langle (K^*K + \alpha L)(x_n - x_0), x_n - x_0 \rangle - \langle K^*(y_n - y_0), x_n - x_0 \rangle \\ &\geq \alpha c_0 \|x_n - x_0\|^2 - \|K^*(y_n - y_0)\| \|x_n - x_0\|, \end{aligned}$$

där den sista olikheten följer av (26) och av att $\langle K^*Kx, x \rangle = \langle Kx, Kx \rangle$. Följaktligen är

$$\|x_n - x_0\| \leq \frac{1}{\alpha c_0} \|K^*(y_n - y_0)\|,$$

så vi ser att $x_n \to x_0$, då $y_n \to y_0$. □

Sats 2.3.9. *Avbildningen R_α är linjär och kan skrivas på den explicita formen*

$$R_\alpha(y) = x,$$

med

$$R_\alpha = (K^*K + \alpha L)^{-1} K^*. \tag{29}$$

Bevis. Minimipunkten till $G_\alpha(x)$ för ett fixt $y \in Y$ ges av ekvationen

$$\nabla G_\alpha(x) = 0,$$

vilket enligt (28) är ekvivalent med

$$(K^*K + \alpha L)x - K^*y = 0.$$

Detta ger lösningen

$$x^\alpha = R_\alpha(y) = (K^*K + \alpha L)^{-1} K^* y.$$

Låt oss försäkra oss om att operatorn $K^*K + \alpha L$ är injektiv. Antag att $\alpha Lx + K^*Kx = 0$. Skalärmultiplikation med x ger att

$$\langle \alpha Lx + K^*Kx, x \rangle = \alpha \langle Lx, x \rangle + \langle Kx, Kx \rangle = 0,$$

varav följer att $x = 0$.

Att R_α är linjär framgår av att vi i högerledet av (29) har en sammansättning av idel linjära avbildningar. □

2.4 Landweberiteration

Låt $K : X \longrightarrow Y$ vara linjär och sätt

$$F(x, y) = (I - aK^*K)x + aK^*y, \quad a > 0.$$

För ett givet $y \in Y$ definierar vi nu rekursivt en följd $(x_n)_0^\infty$ i X, genom iteration av ekvationen $x = F(x, y)$:

$$\begin{aligned} x_0 &= 0 \\ x_n &= F(x_{n-1}, y), \quad n = 1, 2, 3, \ldots. \end{aligned} \tag{30}$$

Sats 2.4.1. *Följden* (x_n) *kan skrivas på explicit form* $x_n = R_n y$, $n = 0, 1, 2, \ldots$, *där*

$$\begin{aligned} R_0 y &= 0 \\ R_n &= a\sum_{k=0}^{n-1}(I - aK^*K)^k K^*, \quad n = 1, 2, 3, \ldots. \end{aligned} \tag{31}$$

Bevis. Genom induktion över $n \geq 1$. Det är klart att påståendet är sant för $n = 1$. Vi antar att påståendet är sant för godtyckligt n, alltså att $x_n = R_n y$, och betraktar elementet x_{n+1} i följden:

$$\begin{aligned} x_{n+1} &= (I - aK^*K)x_n + aK^*y \\ &= (I - aK^*K)\left(a\sum_{k=0}^{n-1}(I - aK^*K)^k K^*\right)y + aK^*y \\ &= a\sum_{k=0}^{n-1}(I - aK^*K)^{k+1}K^*y + aK^*y \\ &= a\sum_{k=0}^{n}(I - aK^*K)^k K^*y. \end{aligned}$$

□

Sats 2.4.2.

(i) Låt $0 < a < 1/\|K\|^2$. *Familjen* $\{R_n : Y \longrightarrow X\}_{n\in\mathbb{Z}^+}$ *av linjära avbildningar utgör en regulariseringsstrategi, med den diskreta regulariseringsparametern* $\alpha = 1/n$ *och* $\|R_n\| \leq \sqrt{an}$. *Den regulariserade lösningen* $x^{\alpha,\delta} := x^{n,\delta} = R_n y^\delta$ *beräknas med hjälp av iterationen (30), så att*

$$\begin{aligned} x^{0,\delta} &= 0 \\ x^{n,\delta} &= (I - aK^*K)x^{n-1,\delta} + aK^*y^\delta, \quad n = 1, 2, 3, \ldots. \end{aligned}$$

Ett val $n = n(\delta)$ *är tillåtet om* $n(\delta) \to \infty$ *och* $\delta^2 n(\delta) \to 0$ *då* $\delta \to 0$.

(ii) Låt $x = K^*z$, *för något* $z \in Y$, *med* $\|z\| \leq E$ *och låt* $0 < c_1 < c_2$, *där* $c_1,\ c_2 \in \mathbb{R}$. *För varje val av strategi* $n = n(\delta)$ *med*

$$c_1 E/\delta \leq n(\delta) \leq c_2 E/\delta,$$

gäller feluppskattningen

$$\|x^{n(\delta),\delta} - x\| \leq c_3\sqrt{\delta E},$$

för något c_3 *beroende på* c_1, c_2 *och* a.

(iii) *Låt $x = K^*Kz$, för något $z \in X$, med $\|z\| \leq E$ och låt $0 < c_1 < c_2$. För varje val av strategi $n = n(\delta)$ med*

$$c_1(E/\delta)^{2/3} \leq n(\delta) \leq c_2(E/\delta)^{2/3},$$

gäller feluppskattningen

$$\|x^{n(\delta),\delta} - x\| \leq c_3 E^{1/3}\delta^{2/3},$$

för något c_3 beroende på c_1, c_2 och a.

Bevis. (i): Låt $(\mu_j, x_j, y_j)_{j\in J}$ vara ett singulärvärdessystem för K. Vi ska visa att $R_n y$ har representationen

$$R_n y = a\sum_{j\in J}\mu_j\sum_{k=0}^{n-1}(1-a\mu_j^2)^k\langle y, y_j\rangle x_j, \quad n = 1, 2, 3, \ldots \tag{32}$$

Detta visar vi återigen genom induktion över $n \geq 1$. För $n = 1$ får vi att

$$a\sum_{j\in J}\mu_j\sum_{k=0}^{0}(1-a\mu_j^2)^k\langle y, y_j\rangle x_j = a\sum_{j\in J}\mu_j\langle y, y_j\rangle x_j = aK^*y.$$

Antag nu att (32) är sant för godtyckligt n. Då har vi att

$$\begin{aligned}
R_{n+1}y &= (I - aK^*K)x^n + aK^*y \\
&= [I - aK^*K]\left(a\sum_{j\in J}\mu_j\sum_{k=0}^{n-1}(1-a\mu_j^2)^k\langle y, y_j\rangle x_j\right) + aK^*y \\
&= a\sum_{j\in J}\mu_j\sum_{k=0}^{n-1}(1-a\mu_j^2)^k\langle y, y_j\rangle(I - aK^*K)x_j + aK^*y \\
&= a\sum_{j\in J}\mu_j\sum_{k=0}^{n-1}(1-a\mu_j^2)^{k+1}\langle y, y_j\rangle x_j + aK^*y \\
&= a\sum_{j\in J}\mu_j\sum_{k=0}^{n}(1-a\mu_j^2)^k\langle y, y_j\rangle x_j.
\end{aligned}$$

Alltså gäller (32) även för $n + 1$. Vidare har vi att

$$a\sum_{j\in J}\mu_j\sum_{k=0}^{n-1}(1-a\mu_j^2)^k\langle y, y_j\rangle x_j = \sum_{j\in J}\frac{1}{\mu_j}(1-(1-a\mu_j^2)^n)\langle y, y_j\rangle x_j,$$

eftersom den inre summan i vänsterledet är geometrisk:

$$\sum_{k=0}^{n-1}(1-a\mu_j^2)^k = \frac{1-(1-a\mu_j^2)^{(n-1)+1}}{1-(1-a\mu_j^2)} = \frac{1-(1-a\mu_j^2)^n}{a\mu_j^2}.$$

Låter vi alltså

$$q(n,\mu) = 1 - (1 - a\mu^2)^n,$$

så kan vi skriva R_n på formen

$$R_n y = \sum_{j \in J} \frac{q(n,\mu_j)}{\mu_j} \langle y, y_j \rangle x_j,$$

för alla $y \in Y$. Enligt sats 2.1.4 del (ii) är q ett regulariserande filter för K, om vi gör substitutionen $\alpha = 1/n$. Del (i) följer alltså av sats 2.1.1 och sats 2.1.4.

(ii): Om vi kombinerar den allmänna feluppskattningen (8) med sats 2.1.2 och sats 2.1.4 får vi feluppskattningen

$$\begin{aligned}
\|x^{n,\delta} - x\| &\leq \delta\sqrt{an} + \frac{1}{\sqrt{2a}}\frac{1}{\sqrt{n}}\|z\| \\
&\leq \delta\sqrt{ac_2 E/\delta} + \frac{1}{\sqrt{2a}}\frac{1}{\sqrt{c_1 E/\delta}}E \\
&= (\sqrt{ac_2} + \frac{1}{\sqrt{2ac_1}})\sqrt{\delta E} = c_3\sqrt{\delta E},
\end{aligned}$$

där vi ser att c_3 enbart beror på c_1, c_2 och a. Den sista olikheten erhålles genom att i första termen ersätta n med den övre begränsningen för $n = n(\delta)$ och i andra termen med den undre begränsningen.

(iii): Helt i analogi med del (ii) har vi mutatis mutandis att

$$\begin{aligned}
\|x^{n,\delta} - x\| &\leq \delta\sqrt{an} + \frac{1}{a}\frac{1}{n}\|z\| \\
&\leq \delta\sqrt{ac_2(E/\delta)^{2/3}} + \frac{1}{a}\frac{1}{c_1(E/\delta)^{2/3}}E \\
&= (\sqrt{ac_2} + \frac{1}{ac_1})E^{1/3}\delta^{2/3} = c_3 E^{1/3}\delta^{2/3},
\end{aligned}$$

där vi ser att c_3 enbart beror på c_1, c_2 och a. □

3 Exempel på inversa problem i tillämpningar; numeriska experiment med regulariseringsmetoder.

Vi kommer att laborera med tre exempelproblem: det en- respektive tvådimensionella avfaltningsproblemet samt tomografiproblemet. Dessa ges ursprungligen i en kontinuerlig tappning, men diskretiseras på lämpligt sätt för beräkningssyften.

3.1 Avfaltningsproblemet i en dimension

Definition 3.1.1. *Låt X vara ett metriskt rum och låt $f : X \longrightarrow \mathbb{R}$. Det slutna höljet av mängden*

$$\{x \in X : f(x) \neq 0\}$$

kallas för stödet *eller* stödmängden *till funktionen f, och betecknas* $\operatorname{supp}(f)$. *Vi säger att f har* ändligt -, begränsat- *respektive* kompakt stöd *om* $\operatorname{supp}(f)$ *är ändlig, begränsad respektive kompakt.*

Om X är ett euklidiskt rum innebär kompakt stöd detsamma som begränsat stöd. Detta är alltså ett i matematiken brukligt sätt att säga att funktionen f är lika med noll utanför något klot kring noll.

Definition 3.1.2. *En* absolut integrerbar funktion *eller* L^1-funktion $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *är sådan att normen*

$$\|f\|_1 = \int_{\mathbb{R}^n} |f(x)|dx < \infty.$$

Det linjära rummet av absolut integrerbara funktioner betecknas $L^1(\mathbb{R}^n)$. *En* lokalt integrerbar funktion *är sådan att integralen*

$$\int_{\|x\|<r} |f(x)|dx$$

är definierad och ändlig för varje $r > 0$.

Anmärkning. En absolut integrerbar funktion f kan med godtycklig noggrannhet approximeras av en absolut integrerbar funktion med kompakt stöd. Vi kan ge detta följande mer precisa formulering: för varje $\epsilon > 0$ finns en absolut integrerbar funktion f_0 med kompakt stöd som är sådan att

$$\|f - f_0\|_\infty < \epsilon,$$

där $\|_\|_\infty$ är maximumnormen på $L^1(\mathbb{R}^n)$,

$$\|f\| := \max_{x \in \mathbb{R}^n} \|f(x)\|.$$

Låt nämligen $D \subset \mathbb{R}^n$ vara en ändlig sluten kub sådan att $\|f(x)\| < \epsilon$ för alla $x \notin D$. En lämplig approximerande funktion med kompakt stöd är då

$$f_0(x) = f_D(x) = \begin{cases} f(x) & \text{om } x \in D, \\ 0 & \text{annars.} \end{cases} \tag{33}$$

Definition 3.1.3. *Låt $f,\ g : \mathbb{R}^n \longrightarrow \mathbb{R}$ vara absolut integrerbara. Då definieras* faltningen *(från tyskans* faltung*, vikning) eller* faltningsprodukten *eller* konvolutionen *$f * g$ av f och g enligt ekvationen*

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy.$$

Med variabelsubstitutionen $z = x - y$ ser vi i fallet $n = 1$ att

$$\int_{-\infty}^{\infty} f(x-y)g(y)dy = -\int_{\infty}^{-\infty} g(x-z)f(z)dz = \int_{-\infty}^{\infty} g(x-z)f(z)dz,$$

så att faltningen är en kommutativ operation, vilket även är fallet i flera dimensioner. Det visar sig att faltningsprodukten är kommutativ, associativ och distributiv. Faltningsprodukten dyker upp i många tillämpningar och i flera grenar av matematiken, bl.a. i linjär analys, sannolikhetslära och matematisk statistik. Exempelvis är sannolikhetsfördelningen för summan av två oberoende stokastiska variabler lika med faltningen av de två variablernas respektive fördelningar. Ett oskarpt fotografi kan representeras som faltningen av den skarpa bilden med formen på kameralinsens bländare.

Det till detta och nästa avsnitt hörande inversa problem är följande:

> **Avfaltningsproblemet:** Låt $k : \mathbb{R}^n \longrightarrow \mathbb{R}$ vara absolut integrerbar. Definiera den linjära operatorn $K : L^1(\mathbb{R}^n) \longrightarrow L^1(\mathbb{R}^n)$ enligt
>
> $$Kf = k * f.$$
>
> Låt nu $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ vara absolut integrerbar. Problemet består i att lösa ekvationen
>
> $$Kf = g.$$

Funktionen k kallas i detta sammanhang för *faltningskärnan* (för operatorn K). Vi inskränker vår undersökning till fallen $n = 1$ och $n = 2$, det en- respektive två-dimensionella avfaltningsproblemet. Vi kommer dock inte att närmare arbeta med problemet i denna ursprungliga kontinuerliga tappning, utan istället med en diskretiserad version av det. Detta eftersom vi å ena sidan i denna uppsats inte studerar regulariseringsteori för oändligtdimensionella Hilbertrum, utan endast specialfallet med euklidiska rum. Å andra sidan är vi i beräkningssammanhang alltid tvungna att göra en diskret representation av problemen, så för praktiska syften gör det ingen skillnad.

En lämplig diskretisering av faltningsprodukten i en dimension är som följer: vi gör en indelning av hela den reella tallinjen i jämnt fördelade punkter med avståndet eller steglängden Δt emellan. Vi väljer ett antal av dessa punkter $p_1, \ldots, p_l$, med $p_i = p_1 + (i-1)\Delta t,\ i = 1, \ldots, l$, i något lämpligt "stödintervall" för faltningskärnan $k(t)$, alltså ett intervall utanför vilket funktionen ligger nära noll. Därefter bildar vi en vektor k i $\mathbb{R}^l$ av funktionsvärdena i respektive punkter:

$$k = \begin{bmatrix} k(p_1) \\ \vdots \\ k(p_l) \end{bmatrix} = \begin{bmatrix} k_1 \\ \vdots \\ k_l \end{bmatrix}.$$

På samma sätt väljer vi ett antal av dessa punkter $q_1, \ldots, q_m$ i ett lämpligt "stödintervall" för $x(t)$ och bildar vektorn

$$x = \begin{bmatrix} x(q_1) \\ \vdots \\ x(q_m) \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}.$$

Uppgiften består nu i att finna en för faltningen $y(t) = k * x(t)$ representerande vektor $y \in \mathbb{R}^n$, vilket innebär att välja en lämplig dimension n på vektorn y och att uttrycka elementet y_i i termer av elementen i k och x, vilket får fungera som approximation av det exakta värdet på faltningen $y(r_i)$ i punkten r_i.

För resonemangets skull ska vi utvidga dessa tre vektorer till tre oändliga följder (k_i), (x_i) och (y_i) i $\mathbb{R}$, där $i \in \mathbb{Z}$, som enbart består av nollor så att säga utanför de ändliga vektorerna k, x respektive y, alltså med

$$\begin{aligned} k_i &= 0, \quad \text{om } i \notin \{1, \ldots, l\}, \\ x_i &= 0, \quad \text{om } i \notin \{1, \ldots, m\} \text{ och} \\ y_i &= 0, \quad \text{om } i \notin \{1, \ldots, n\}. \end{aligned}$$

Dessa följder svarar mot punktföljderna (p_i), (q_i) respektive (r_i) på den reella tallinjen, där

$$p_i = p_1 + (i-1)\Delta t, \quad q_i = q_1 + (i-1)\Delta t, \quad r_i = r_1 + (i-1)\Delta t,$$

för $i \in \mathbb{Z}$. Värdemängderna för dessa punktföljder sammanfaller och är lika med mängden av punkter i vår på förhand valda indelning av $\mathbb{R}$.

Notera att det vi har gjort också kunde beskrivas som att vi först approximerar $x(t)$ och $k(t)$ med funktioner med kompakt stöd på samma sätt som i (33) och därefter utför en diskretisering av dessa approximationer över $\mathbb{R}$.

Vi gör nu följande approximation av $y(r_i)$:

$$y(r_i) = k * x(r_i) = \int_{-\infty}^{\infty} k(t)x(r_i - t)dt$$
$$\approx \sum_{j=-\infty}^{\infty} k(p_j)x(r_i - p_j)\Delta t. \tag{34}$$

En term i denna summa är nollskild endast om $j \in \{1, \ldots, l\}$ och $r_i - p_j = q_h$, för något $h \in \{1, \ldots, m\}$, alltså endast om

$$r_i = p_j + q_h = (p_1 + q_1) + (j + h - 2)\Delta t, \quad 1 \leq j \leq l,\ 1 \leq h \leq m.$$

Detta ger oss

$$n = l + m - 1$$

lämpliga punkter för vektorn y med valet

$$r_1 = p_1 + q_1 \text{ och } r_i = r_1 + (i-1)\Delta t,\ i = 1, \ldots, n.$$

För $i,\ j \in \mathbb{N}$ blir då

$$\begin{aligned} r_i - p_j &= (p_1 + q_1) + (i-1)\Delta t - (p_1 + (j-1)\Delta t) \\ &= q_1 + (i-j)\Delta t = q_{i-j+1}. \end{aligned}$$

Vi får alltså för elementet y_i i vektorn y att

$$y_i = \sum_{j=-\infty}^{\infty} k(p_j)x(r_i - p_j)\Delta t = \sum_{j=-\infty}^{\infty} k(p_j)x(q_{i-j+1})\Delta t$$
$$= \Delta t \sum_{j=-\infty}^{\infty} k_j x_{i-j+1} = \Delta t \sum_{\substack{1 \leq j \leq l \\ 1 \leq i-j+1 \leq m}} k_j x_{i-j+1}. \tag{35}$$

Låt nu avbildningen $K : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ vara sådan att det i :te elementet i $y = Kx$, för ett givet $x \in \mathbb{R}^m$, ges av ekvation (35). K är alltså en diskret approximation av faltningsprodukten. Vi ska visa att K är en linjär avbildning och bestämma matrisen för avbildningen med avseende på standardbaserna för $\mathbb{R}^m$ och $\mathbb{R}^n$.

Vi kan göra detta genom att anta att K är linjär, se vad matrisen i så fall måste vara och därefter helt enkelt konstatera att denna matris faktiskt ger en matrisframställning av avbildningen K, varav förstås följer att K är linjär. Antag alltså att K är linjär. Låter vi alla inblandade vektorer vara på kolonnform är A av typ $n \times m$. Låt $e = (e^1, \ldots, e^m)$ vara standardbasen för $\mathbb{R}^m$ och sätt

$$y^\lambda = Ke^\lambda,$$

för $\lambda = 1, \ldots, m$. I standardbasen har vi att $[y^\lambda] = y^\lambda$, varför kolonn nummer λ i matrisen $A = [K]_{st.baser}$ måste vara lika med y^λ. Vi får för $i = 1, \ldots, n$ att

$$y_i^\lambda = \Delta t \sum_{\substack{1 \le j \le l \\ 1 \le i-j+1 \le m}} k_j e_{i-j+1}^\lambda = \begin{cases} \Delta t \cdot k_{i-\lambda+1} & \text{om } 1 \le i - \lambda + 1 \le l, \\ 0 & \text{annars,} \end{cases}$$

eftersom $i - j + 1 = \lambda$ omm $j = i - \lambda + 1$. Av detta följer att matrisen

$$A = \Delta t \cdot B(k, m) = \Delta t \cdot \begin{bmatrix} k_1 & 0 & \ldots & 0 \\ \vdots & k_1 & \ldots & 0 \\ k_l & \vdots & & \vdots \\ 0 & k_l & \ldots & k_1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & k_l \end{bmatrix}.$$

Denna matris uppfyller villkoret att

$$A[x]_{\text{st.bas}} = [Kx]_{\text{st.bas}},$$

för alla $x \in \mathbb{R}^m$, varav följer att K är en linjär avbildning. I de numeriska experimenten utelämnar vi för enkelhetens skull skalfaktorn Δt och använder oss direkt av matrisen $B(k, m)$.

Vi skriver två funktioner i *MatLab*, disfalt(k, m) och disfunk(a, b, n, f), där den första genererar faltningsmatrisen $B(k, m)$ givet vektorn k och dimensionen m för definitionsrummet till den diskretiserade faltningsoperatorn, alltså storleken på lösningsvektorn x. Den andra diskretiserar funktionen $f(t)$ över intervallet $[a, b]$ i n punkter och returnerar kolonnvektorn f av typen $n \times 1$. Högerledet y beräknas därefter som

$$y = \text{disfalt}(k, m) \cdot \text{disfunk}(a, b, m, x),$$

där $k = \text{disfunk}(c, d, l, k)$. Vi sätter det perturberade högerledet y^δ till

$$y^\delta = y + \epsilon \cdot \text{randn}(\text{size}(y)), \tag{36}$$

där *MatLab*-funktionen randn ger en slumpmatris av given storlek med värden normalfördelade kring 0 med standardavvikelsen 1, alltså den standardiserade normalfördelningen. Ett godtyckligt element i slumpvektorn är mindre än eller lika med 2 med en sannolikhet på 98%, vilket betyder att ett godtyckligt element i y^δ är på avståndet högst 2ϵ från motsvarande element i y med samma sannolikhet.

3.1.1 Numeriska experiment med avfaltningsproblemet i en dimension

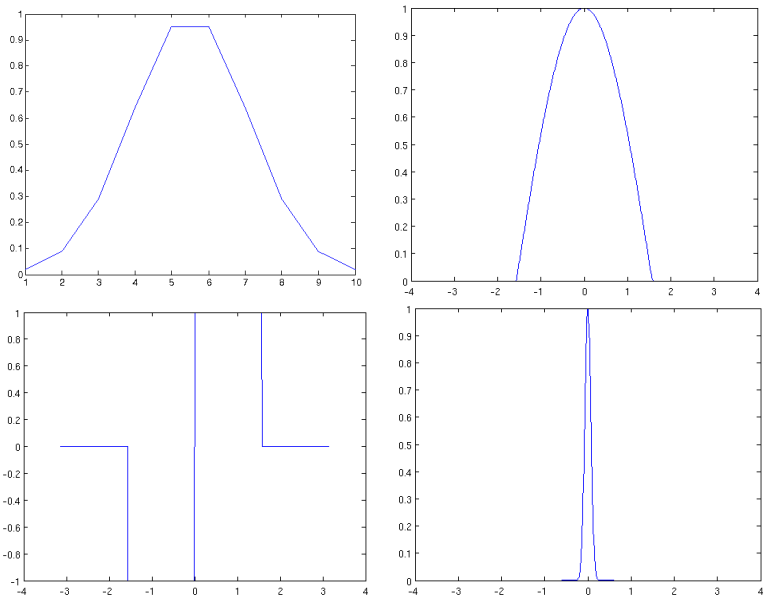
Låt oss som faltningskärna använda en typisk ”spikfunktion”:

$$k(t) = e^{-ct^2},$$

där c är en konstant. Till exakta lösningar väljer vi tre funktioner, en med slät och enkel kurva, en med en hoppdiskontinuitet och en med en ganska spetsig ”spik”:

$$\begin{aligned} x_1(t) &= \cos t \\ x_2(t) &= \operatorname{sgn} t \\ x_3(t) &= e^{-100t^2}, \end{aligned}$$

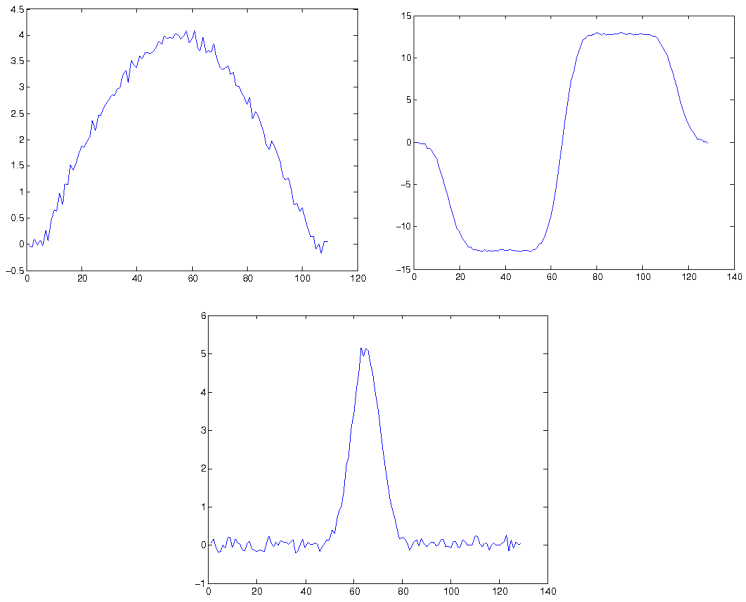
där vi kan anta att dessa ekvationer är uppfyllda för t i intervallet $[-\pi/2, \pi/2]$, men att $x_1(t)$, $x_2(t)$ och $x_3(t)$ är lika med noll utanför detta intervall, ett enkelt sätt att erhålla funktioner med kompakt stöd. Dessa tre funktioner och faltningskärnan visas i figur 1. Vi sätter $\epsilon = 0.1$ i ekvation (36). Detta ger det perturberade högerledet y^δ såsom visas i figur 2.



Figur 1: Från vänster till höger: den diskretiserade faltningskärnan $k(t)$ samt de tre funktionerna x_1, x_2 respektive x_3.

Det kunde vara intressant att se vad som händer om vi inte utnyttjar någon typ av regulariseringsmetod här, utan istället direkt löser ekvationssystemet

$$Kx = y^\delta,$$



Figur 2: Från vänster till höger: perturberade högerled för x_1, x_2 respektive x_3, alltså faltningen av dessa funktioner plus en slumpterm.

vilket i *MatLab* kan göras genom att beräkna uttrycket

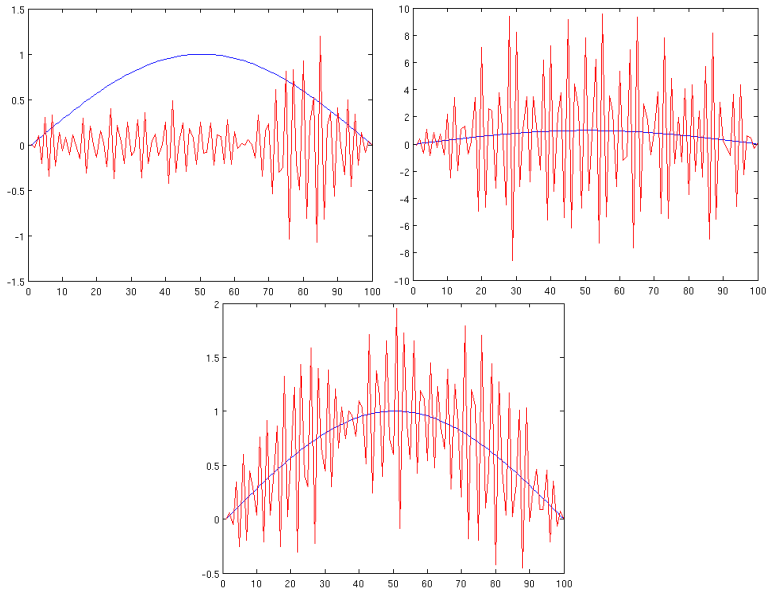
$$z = \text{disfalt}(k, m) \backslash y^{\delta}.$$

Resultatet för varierande värden på ϵ visas i figur 3 och det demonstrerar tydligt behovet och användbarheten av regulariseringsteori.

Tikhonovregularisering. Vi diskretiserar $k(t) = e^{-t^2}$ över intervallet $[-2, 2]$ i 10 punkter. För alla tre fallen diskretiserar vi $x(t)$ över intervallet $[-\pi/2, \pi/2]$ med $m = 100$. Vi vill kunna uppfatta det så att samma steglängd Δt utnyttjas överallt, vilket är möjligt genom en "omskalning" av funktionen $k(t)$. Denna "omskalning" bestämmes av konstanten c: diskretiseringen av $k(t) = e^{-t^2}$ i 10 punkter över intervallet $[-2, 2]$ är ekvivalent med en diskretisering av funktionen $k(t) = e^{-ct^2}$ i 10 punkter över ett mindre intervall $[a', b']$ som ger steglängden $\Delta t = \pi/99$, för några c, a' och b'. I detta fall får vi

$$c = \frac{4 \cdot 22^2}{\pi^2}, \; a' = -\frac{\pi}{22} \text{ och } b' = \frac{\pi}{22}.$$

För detta särskilda problem kan vi utan vidare utföra Tikhonovregulariseringen genom direkt utnyttjande av formel (19). Prövar vi oss fram med några olika värden på regulariseringsparametern α, med den exakta lösningen x_1, får vi resultatet i figur 4. Dessa grafer utgör en illustration av den idé som framföres i avsnitt 2.1, att regulariseringen i praktiken kan beskrivas som att träffa ett val mellan stabilitet i lösningen och närhet till den



Figur 3: Från vänster till höger: försök att bestämma x genom att direkt lösa ekvationssystemet för ϵ =0.1, 0.01 respektive 0.001. I den övre vänstra bilden har vi delat z med 100 därför att elementen fluktuerar mellan -120 och 120.

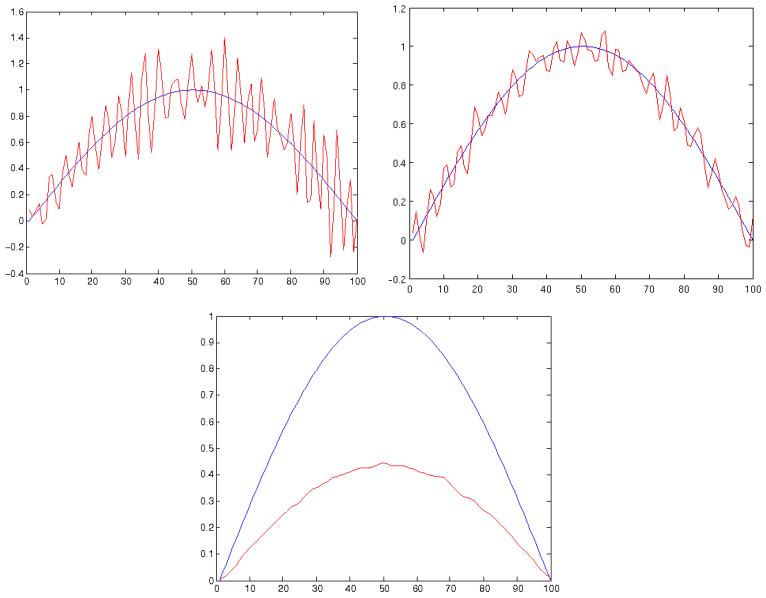
exakta lösningen, eller till verkligheten. Här ser vi tydligt att med $\alpha = 0.01$ får vi en lösning som i medel ansluter nära till x_1, men som är ostadig. Med $\alpha = 10$ får vi istället en jämnare och stabilare lösning, men som i medel ligger betydligt längre bort från x_1. De tre bilderna antyder alltså att ett optimalt val av α ligger i intervallet $[0.01, 10]$ för lösningen x_1.

Vi kan utföra optimeringen för tre olika metriker på $\mathbb{R}^m$: det euklidiska avståndet $\|a-b\| = \|a-b\|_2$, den av maximumnormen inducerade $\|a-b\|_\infty$ och medelvärdet av absolutbeloppet på differensen

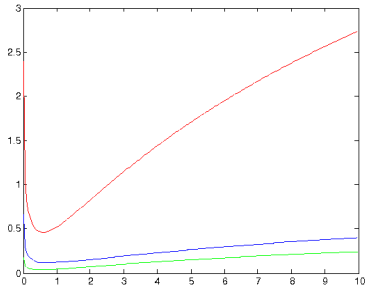
$$M(a,b) := \frac{1}{m}\sum_{i=1}^{m} |a_i - b_i| = \frac{1}{m}\|a-b\|_1.$$

Om vi ritar kurvorna till $\|x_1^\alpha - x_1\|$ för de tre olika metrikerna med $\alpha \in [0.01, 10]$ som oberoende variabel får vi resultatet i figur 5. Dessa kurvor illustrerar återigen den typiska situationen eller typfallet i samband med regularisering, det att välja α kan beskrivas som att försöka finna jämviktspunkten. Såsom kurvorna visar finns här för varje särskild metrik en entydigt bestämd minimipunkt. Denna minimipunkt α^* utgör då det optimala valet av parameter under en viss norm eller metrik. De tre olika metrikerna pekar mot ungefär samma optimala val för en given vektor x $(= x_1,\ x_2,\ x_3)$.

Den euklidiska metriken antyder att ett optimalt val för $x = x_1$ torde

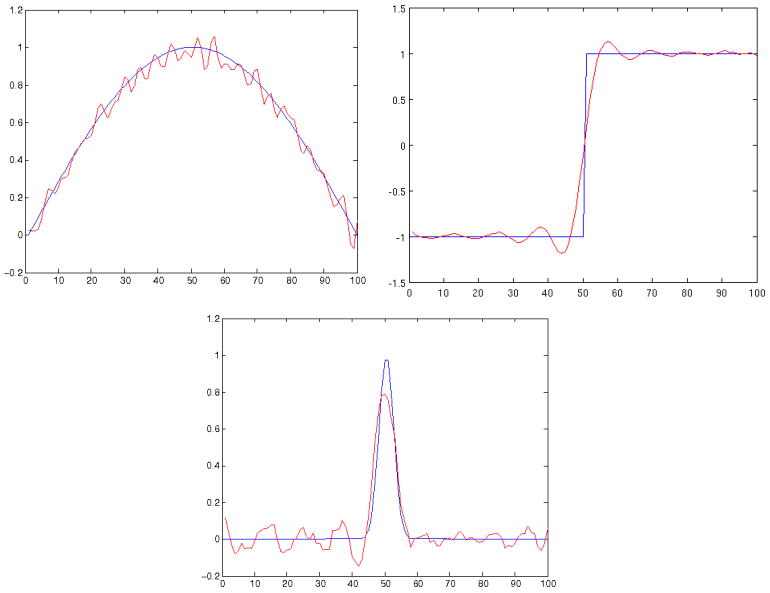


Figur 4: Från vänster till höger: Tikhonovregularisering för $\alpha = 0.01$, $\alpha = 0.1$ och $\alpha = 10$ med den exakta lösningen $x = x_1$.



Figur 5: Optimering av α under tre olika metriker på $\mathbb{R}^n$. På x-axeln har vi $\alpha \in [0.01, 10]$ och på y-axeln $\|x - x^{\alpha,\delta}\|$. Den översta kurvan visar resultatet för den euklidiska normen och den understa kurvan för maximumnormen.

ligga någonstans kring $\alpha = 0.5$, för x_2 kring $\alpha = 0.3$ och för x_3 kring $\alpha = 10$. Vi ska dock inte glömma att vi har infört en slumpterm i y^δ som gör att α^* varierar något från beräkning till beräkning fastän x hålles fixt.



Figur 6: Från vänster till höger: Tikhonovregularisering med optimalt val av α, enligt den euklidiska normen, för de exakta lösningarna x_1, x_2 respektive x_3.

Generaliserad Tikhonovregularisering. Vi sätter

$$L = B^*B,$$

där $B : \mathbb{R}^m \longrightarrow \mathbb{R}^{m-1}$ är den *diskreta derivatan* eller *derivationsoperatorn*

$$Bx = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ \vdots \\ x_m - x_{m-1} \end{bmatrix}.$$

Om vi tänker oss att x representerar en funktion $x(t)$ kan vi uppfatta Bx som en approximation av en diskretisering av $x'(t)$, med

$$Bx \approx \begin{bmatrix} x'(t_1) \\ x'(t_2) \\ \vdots \\ x'(t_{m-1}) \end{bmatrix}.$$

Matrisen för B i standardbaser är av typ $(m-1) \times m$ och

$$[B]_{ij} = \begin{cases} -1 & \text{om } i = j, \\ 1 & \text{om } i+1 = j, \\ 0 & \text{annars,} \end{cases}$$

för $1 \leq i \leq m-1,\ 1 \leq j \leq m$. Detta gör det lätt att skriva en *MatLab*-funktion matrisL(m) som genererar $m \times m$-matrisen för operatorn L. Med detta så kan mutatis mutandis samma funktioner utnyttjas som för Tikhonovregulariseringen.

Det är rimligt att förvänta sig att resultatet av den generaliserade Tikhonovregulariseringen med detta val av B ger en slätare kurva jämfört med vanlig Tikhonovregularisering, vilket för en slät kurva såsom $x = x_1$ torde innebära en närmare approximation till x_1 för optimalt val av α. Att detta är rimligt att förvänta sig kan man inse på följande sätt: för ett givet $\alpha > 0$ vill vi minimera uttrycket

$$\|Kx^\alpha - y\|^2 + \alpha\|Bx^\alpha\|^2$$

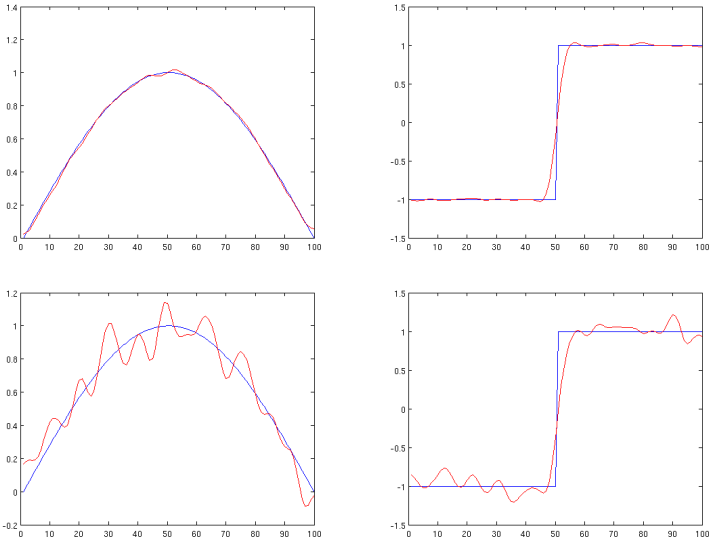
med avseende på x^α för det givna högerledet y. Ju större avståndet är mellan på varandra följande punkter i x^α, desto större kommer $\|Bx^\alpha\|$ att bli. Detta innebär alltså att förhållandet mellan punkterna i x^α här är av relevans. I fallet med vanlig Tikhonovregularisering betraktar vi istället $\|x^\alpha\|$, vars värde inte påverkas av relationen mellan punkterna i x^α — de är så att säga oberoende av varandra.

Med $\alpha = 25$ får vi för $x = x_1$ resultatet i figur 7. Som vi ser är den optimala regulariserade lösningen här överlag snäppet bättre jämfört med figur 6. Vi har också prövat att variera felet ϵ i y^δ. Resultatet för $x = x_3$ visas inte i figuren därför att det inte bidrar med något nytt av principiellt intresse för vår förståelse av regulariseringsmetoder.

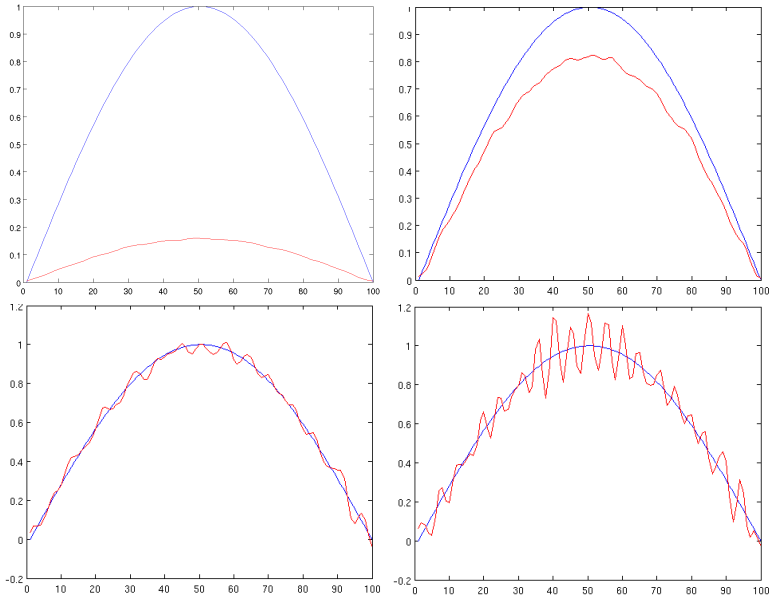
Landweberiteration. Villkoret på parametern för att iterationen ska fungera är enligt sats 2.4.2 att $0 < a \leq 1/\|K\|^2$. Vi har här på prov satt $a = 0.1$, vilket tycks fungera. Vi skriver en *MatLab*-funktion Lw(p, a) som returnerar den p:te Landweberiterationen $x^p \in \mathbb{R}^m$ för ett visst val av värde på parametern a. Även i detta fall är alltså regulariseringen rakt på sak, eftersom vi utan omvägar kan utnyttja matrisen för den diskreta faltningsoperatorn och därmed regulariseringsformlerna direkt. Figur 8 antyder att ett optimalt val av antal iterationer torde ligga någonstans nära $p = 100$, vilket ger värdet på $\alpha = 1/100 = 0.01$.

3.2 Avfaltningsproblemet i två dimensioner

I diskretiseringen av den två-dimensionella faltningsoperatorn K representeras en funktion $x : \mathbb{R}^2 \longrightarrow \mathbb{R}$ av en $m_1 \times m_2$-matris bestående av funktionsvärdena i ett antal jämnt fördelade punkter tillhörande någon "stödrektangel" i $\mathbb{R}^2$, alltså ett rektangulärt område som omfattar supp(x). Vi



Figur 7: Generaliserad Tikhonovregularisering med de exakta lösningarna x_1 respektive x_2. I de övre två bilderna har vi $\epsilon = 0.1$ och i de undre $\epsilon = 1$.



Figur 8: Från vänster till höger: Landweberiteration med $x = x_1$ för ett varierande antal iterationer $p = 1$, 10, 100, respektive 500.

låter som i föregående avsnitt den representerande matrisen ha samma beteckning som funktionen, och låter det av sammanhanget framgå vad som åsyftas. Motsvarigheten till ekvation (34) är rakt på sak:

$$y(r_i, r_j) = k * x(r_i, r_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k(s,t)x(r_i - s, r_j - t)\ dsdt$$
$$\approx \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} k(s_{\tau_1}, t_{\tau_2})x(r_i - s_{\tau_1}, r_j - t_{\tau_2})\Delta t^2.$$

Det är för numeriska beräkningar inte så lämpligt att såsom i det endimensionella fallet försöka utnyttja ett explicit uttryck för matrisen till den diskretiserade faltningsoperatorn i två dimensioner. Detta eftersom vi med relativt få indelningspunkter av relevant område i planet, alltså med få "pixlar" eller bildpunkter, erhåller en otymplig matris. Istället kommer vi helt enkelt använda oss av den i *MatLab* inbyggda funktionen conv2, vars algoritm i princip utnyttjar ovanstående approximationsformel:

$$Kx \approx K(x) := \text{conv2}(x, k),$$

bortsett från en skalfaktor Δt^2. Om matrisen x är av typ $m_1 \times m_2$ och k av typ $l_1 \times l_2$, så är $K(x)$ av typ $n_1 \times n_2 = (m_1 + l_1 - 1) \times (m_2 + l_2 - 1)$ (jmf. föregående avsnitt).

Eftersom vi som sagt inte vill använda matrisen för den diskreta faltningsoperatorn har vi också behov av ett lämpligt *MatLab*-uttryck för adjunkten till K. I *MatLab* är flipud och fliplr funktioner sådana att givet matrisen A är flipud(A) speglingen av A i en horisontell symmetrilinje genom A (så att säga den upp-och nedvända matrisen till A) och fliplr(A) speglingen av A i en vertikal symmetrilinje genom A (så att säga den höger- och vänstervända matrisen till A). Låt exempelvis

$$A = \begin{bmatrix} 1 & 5 \\ 3 & 9 \\ 2 & 7 \end{bmatrix}.$$

Då är

$$\text{flipud}(A) = \begin{bmatrix} 2 & 7 \\ 3 & 9 \\ 1 & 5 \end{bmatrix} \text{ och } \text{fliplr}(A) = \begin{bmatrix} 5 & 1 \\ 9 & 3 \\ 7 & 2 \end{bmatrix}.$$

Låt nu

$$k' = \text{flipud}(\text{fliplr}(k)) = \text{fliplr}(\text{flipud}(k)). \tag{37}$$

Vi ska helt utan vidare bara konstatera att

$$K^* y \approx K^{ad}(y) := \text{conv2}(y, k', \text{'valid'}), \tag{38}$$

bortsett från en skalfaktor. Läsaren kan experimentellt verifiera denna formel i *MatLab* genom att kontrollera att

$$\text{sum}(\text{sum}(x. * K^{ad}(y))) = \text{sum}(\text{sum}(y. * K(x))),$$

för några val av matriser x, y och k. Parametern 'valid' gör så att conv2 i ekvation (38) returnerar en matris av samma typ som x.

Vi har behov av att kunna representera argumenten omväxlande som matriser respektive kolonnvektorer. Detta görs i *MatLab* med funktionen reshape, som definieras på så sätt att om x är av typ $m \times n$, så är reshape$(x, mn, 1)$ den kolonnvektor vi får genom att så att säga stapla kolonnerna i x på varandra med den första överst. Denna vektor betecknas också $x(:)$. Omvänt erhåller vi x som

$$x = \text{reshape}(x(:), m, n).$$

Avbildningen K kan uppfattas omväxlande som en avbildning från matriser respektive kolonnvektorer. Om vi t.ex. talar om matrisen för K är det underförstått att argumenten ska uppfattas som kolonnvektorer av ovanstående typ och inte som matriser.

Liksom i föregående avsnitt kan vi också införa ett godtyckligt stort slumpfel i högerledet genom ekvation (36).

3.2.1 Numeriska experiment med avfaltningsproblemet i två dimensioner

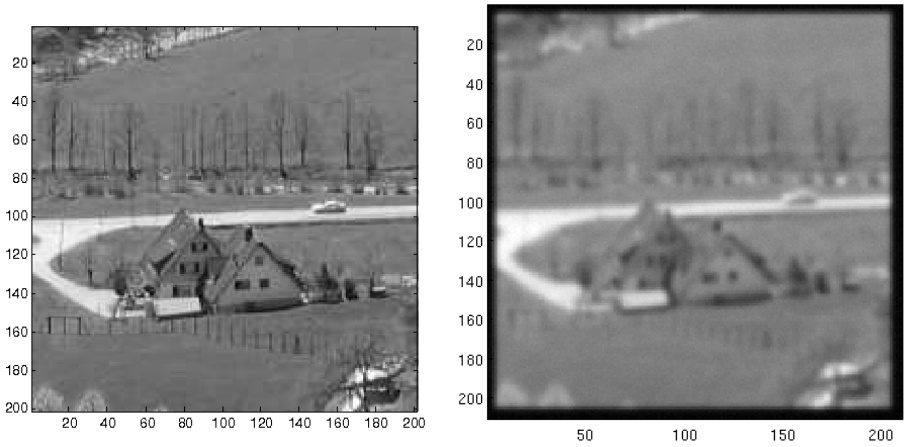
Vi nöjer oss med att studera ett enstaka exempel på avfaltningsproblem i två dimensioner. Vårt X utgöres av en matris av typ 200×200, där elementet x_{ij} representerar ljusstyrkan enligt någon viss gråskala i bildpunkten (i, j) på ett monokromatiskt fotografi, som visas i figur 9. Värdet på x_{ij} varierar mellan 0 och 10^3. Vi kan uppfatta X som diskretiseringen av en reellvärd funktion $X(s, t)$ på $\mathbb{R}^2$ över någon rektangel i planet. Som faltningskärna använder vi oss även här av en typisk "spikfunktion"

$$k(s, t) = e^{-(s^2+t^2)},$$

som diskretiseras över rektangeln $[-2, 2] \times [-2, 2]$ i 10×10 punkter. Resultatet av att falta X med k ger en matris vars element är av storleksordningen 10^4. Därför väljer vi ett ganska stort värde på parametern ϵ för att få ett någorlunda märkbart värde på det relativa felet i det perturberade högerledet Y^δ jämfört med det exakta högerledet Y.

Matriserna ritas i *MatLab* enligt en gråskala med kommandona

$$\text{imagesc}(x);\ \text{axis image};\ \text{colormap(gray)}.$$



Figur 9: Testbilden X och den faltade bilden Y^δ med $\epsilon = 100$.

Tikhonovregularisering. Till skillnad från det endimensionella fallet kan vi inte direkt utnyttja formel (19), emedan detta förutsätter att vi kan beräkna matrisinversen direkt. Istället ska vi begagna oss av *MatLab*-funktionen pcg(A, b), där A är en matris och b en kolonnvektor. Denna funktion försöker lösa ekvationssystemet $Ax = b$ med användande av den s k "preconditioned conjugate gradients method". Förutsättningen för att metoden ska fungera är att A är symmetrisk, och alltså kvadratisk, och positivt definit. Funktionen pcg accepterar som första argument även en linjär avbildning F med

$$Fx = Ax,$$

vilket är användbart för oss då vi nu inte har direkt tillgång till matrisen för K.

Vi sätter

$$F = \alpha I + K^* K.$$

Matrisen för F uppfyller villkoren för pcg. Funktionen F tar som argument en kolonnvektor med mn element. Om vi låter k' vara som i ekvation (37) är pseudokoden för F som följer:

$$\begin{aligned}
&\text{function } u = F(x)\\
&\quad \text{global } m_1,\ m_2,\ k,\ \alpha'\\
&\qquad x = \text{reshape}(x, m_1, m_2)\\
&\qquad z = \text{conv2}(x, k)\\
&\qquad u = \alpha' x + \text{conv2}(z, k', \text{'valid'})\\
&\qquad u = u(:)
\end{aligned}$$

En anmärkning angående pseudkoden. I pseudokoden dyker flera odeklarerade variabler upp som under faktisk implementering i *MatLab* antingen

måste deklareras som globala eller tas med som argument i funktionshuvudet. Emedan valet här står fritt för användaren har vi låtit det vara underförstått att något av detta är gjort för varje odeklarerad variabel. Vissa variabler har vi dock deklarerat som globala redan i pseudokoden, nämligen i det fall författaren funnit det lämpligt att göra så i *MatLab*.

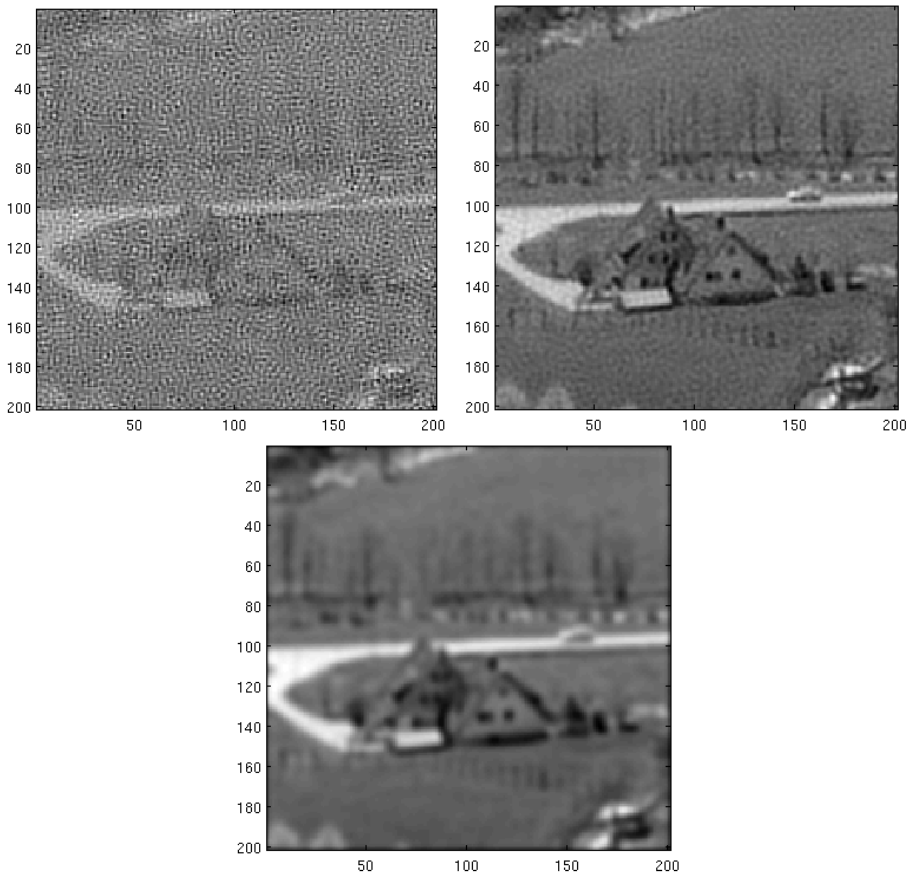
Vi är nu redo att skriva funktionen för tikhonovregulariseringen med pseudokoden

$$
\begin{aligned}
&\text{function } x^\alpha = T(\alpha)\\
&\quad \text{global } m_1, m_2, k, \alpha'\\
&\qquad \alpha' = \alpha\\
&\qquad [n_1, n_2] = \text{size}(y)\\
&\qquad [l_1, l_2] = \text{size}(k)\\
&\qquad m_1 = n_1 - l_1 + 1;\ m_2 = n_2 - l_2 + 1\\
&\qquad z = \text{conv2}(y, k', \text{'valid'})\\
&\qquad x = \text{pcg}(F, z(:))\\
&\qquad x = \text{reshape}(x, m_1, m_2).
\end{aligned}
$$

Istället för att direkt beräkna inversen $(\alpha I + K^*K)^{-1}$ löser vi alltså ekvationssystemet

$$Fx = K^*y.$$

Generaliserad Tikhonovregularisering. I detta sammanhang gör vi bruk av den *diskreta gradienten* eller den "diskreta ∇-operatorn" $B_2 : \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{2mn-m-n}$. Låt oss definiera kolonnvektorn B_2x för en given $m \times n$-matris x så att vi först skriver upp den diskreta derivatan, såsom den definierades i avsnitt 3.1.1, av varje kolonn i x med den första kolonnen överst i vektorn och därefter den diskreta derivatan av varje rad i x med den första raden överst i vektorn. Låt $x^{(j)}$ och $x_{(i)}$ beteckna den $j : te$ kolonnen respektive



Figur 10: Från vänster till höger: Tikhonovregularisering med $\alpha = 0.001$, 1 respektive 100.

den $i:te$ raden i x, $1 \leq i \leq m$, $1 \leq j \leq n$. Detta ger

$$B_2 x = \begin{bmatrix} Bx^{(1)} \\ \vdots \\ Bx^{(n)} \\ Bx^t_{(1)} \\ \vdots \\ Bx^t_{(m)} \end{bmatrix} = \begin{bmatrix} x_{21} - x_{11} \\ \vdots \\ x_{m1} - x_{(m-1)1} \\ \vdots \\ x_{2n} - x_{1n} \\ \vdots \\ x_{mn} - x_{(m-1)n} \\ x_{12} - x_{11} \\ \vdots \\ x_{1n} - x_{1(n-1)} \\ \vdots \\ x_{m2} - x_{m1} \\ \vdots \\ x_{mn} - x_{m(n-1)} \end{bmatrix},$$

där B är den diskreta derivatan.

Tänker vi oss att x representerar en funktion $x(s,t)$ kan vi uppfatta $B_2 x$ som en approximation av en diskretisering av

$$\nabla x = \begin{bmatrix} \frac{\partial x}{\partial s} \\ \frac{\partial x}{\partial t} \end{bmatrix},$$

med

$$B_2 x \approx \begin{bmatrix} \frac{\partial}{\partial s} x(s_1, t_1) \\ \vdots \\ \frac{\partial}{\partial s} x(s_{m-1}, t_n) \\ \frac{\partial}{\partial t} x(s_1, t_1) \\ \vdots \\ \frac{\partial}{\partial t} x(s_m, t_{n-1}) \end{bmatrix}.$$

Vi behöver en metod för beräkning av $L_2 x$ för en given $m \times n$-matris x, där

$$L_2 = B_2^* B_2,$$

som inte kräver att vi har tillgång till matrisen för avbildningen B. Vi kan underlätta det fortsatta resonemanget genom att studera ett exempel. Låt x vara en godtycklig 3×3-matris,

$$x = \begin{bmatrix} x_{11} & - & x_{12} & - & x_{13} \\ | & & | & & | \\ x_{21} & - & x_{22} & - & x_{23} \\ | & & | & & | \\ x_{31} & - & x_{32} & - & x_{33} \end{bmatrix}. \tag{39}$$

Strecken inuti matrisen visar vilka differenser som ryms i B_2x:

$$B_2x = \begin{bmatrix} x_{21} - x_{11} \\ x_{31} - x_{21} \\ x_{22} - x_{12} \\ x_{32} - x_{22} \\ x_{23} - x_{13} \\ x_{33} - x_{23} \\ x_{12} - x_{11} \\ x_{13} - x_{12} \\ x_{22} - x_{21} \\ x_{23} - x_{22} \\ x_{32} - x_{31} \\ x_{33} - x_{32} \end{bmatrix}.$$

Låt oss betrakta matrisen $[B_2]$ för B_2 (i standardbaser). Denna är av typ 12×9. Matrisen x representeras som kolonnvektorn $x(:)$ och vi låter e_{ij} beteckna den basvektor för definitionsrummet $M_{3\times3}(\mathbb{R})$ som har en etta på plats (i, j) och nollor i övriga positioner, vilket svarar mot basvektorn $e_{ij}(:)$ för $\mathbb{R}^9$.

Det finns tre typer av element i en $m \times n$-matris: *hörn-*, *kant-* och *inre* element. I matrisen x finns fyra hörn, fyra kantelement och ett inre element. Låt oss undersöka vad bilden av en basvektor svarande mot ett element av respektive sort blir. Vi får t.ex. att

$$B_2e_{11} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; B_2e_{12} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; B_2e_{22} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

Med $z = L_2x$ får vi att $z = [B_2]^tB_2x$ och

$$z_{ij} = (B_2e_{ij})^tB_2x.$$

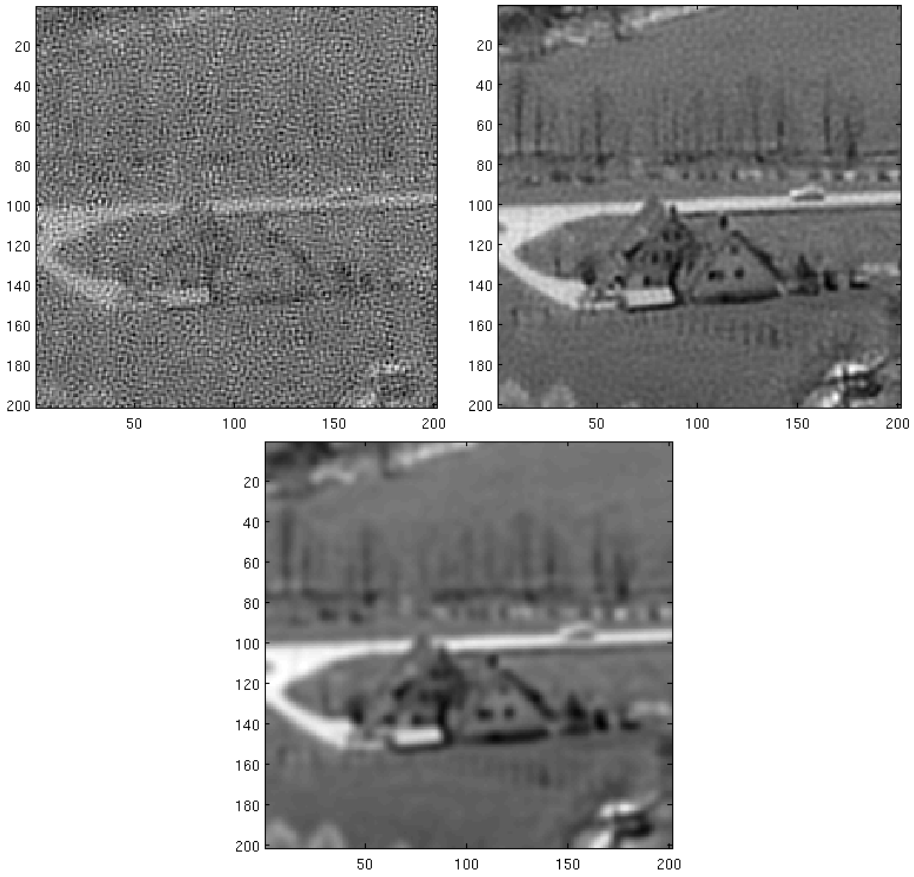
Detta ger att

$$\begin{aligned} z_{11} &= 2x_{11} - x_{21} - x_{12} \quad \text{(hörnelement)}, \\ z_{12} &= 3x_{12} - x_{11} - x_{22} - x_{13} \quad \text{(kantelement)}, \\ z_{22} &= 4x_{22} - x_{21} - x_{32} - x_{23} - x_{12} \quad \text{(inre element)}. \end{aligned}$$

Exemplet illustrerar den allmänna regeln. Definiera en funktion σ på matrisen x sådan att $\sigma(x_{ij})$ är mängden av element i x, bortsett från x_{ij} själv, som tillsammans med x_{ij} ingår i någon differens i B_2x, så att säga mängden av de omgivande elementen i "differensschemat" (39). För en given $m \times n$-matris x beräknas ett godtyckligt element i $z = L_2x$ enligt formeln:

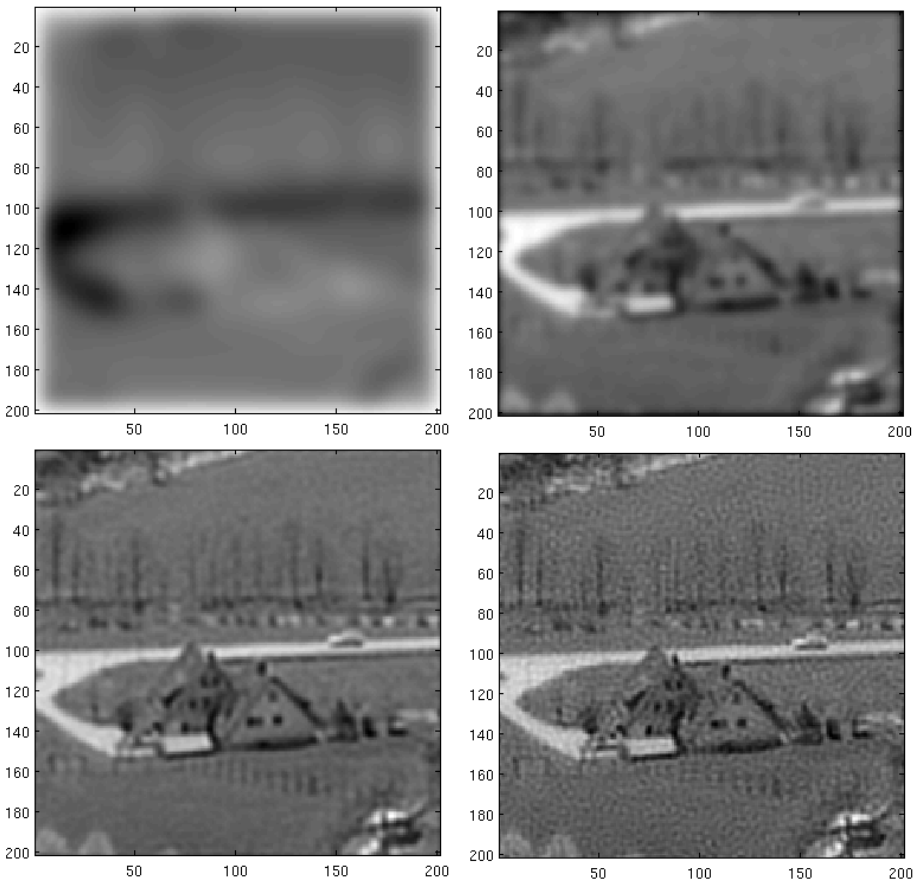
$$z_{ij} = |\sigma(x_{ij})|x_{ij} - \sum_{\tau \in \sigma(x_{ij})} \tau.$$

Detta gör det lätt att skriva en lämplig funktion för L_2 i *MatLab*, varefter vi mutatis mutandis kan genomföra regulariseringen på samma sätt som i föregående avsnitt.



Figur 11: Från vänster till höger: generaliserad Tikhonovregularisering med $\alpha = 0.001$, 1, respektive 100.

Landweberiteration. Här framstår tydligt en del fördelar med landweberiterationen som regulariseringsmetod: den är lättare att implementera i *MatLab* eftersom vi utan vidare kan utnyttja iterationsformeln direkt. Vi



Figur 12: Landweberiteration. Överst till vänster: $a = 0.1$, 10 iterationer. De övriga bilderna visar resultatet för $a = 0.001$ med antal iterationer $p = 1$, 100 respektive 1000.

skriver en *MatLab*-funktion $\mathrm{Lw}(p, a)$, där p är antalet iterationer som ska utföras och a är parametern i sats 2.4.2, med pseudokoden

$$
\begin{aligned}
&\text{function } x = \mathrm{Lw}(p, a)\\
&\qquad m_1 = n_1 - l_1 + 1;\ m_2 = n_2 - l_2 + 1\\
&\qquad x = \mathrm{zeros}(m_1, m_2)\\
&\qquad z_1 = a \cdot \mathrm{conv2}(y, k', \text{'valid'})\\
&\quad \text{for } i = 1 \text{ to } p\\
&\qquad z_2 = \mathrm{conv2}(x, k);\ z_3 = \mathrm{conv2}(z_2, k', \text{'valid'})\\
&\qquad x = x - az_3 + z_1\\
&\quad \text{end.}
\end{aligned}
$$

Vi kan på prov variera parametern a för ett antal iterationer (sätt t.ex. $a = 1,\ 0.1,\ 0.01,\ 0.001,\ 0.0001$). Resultatet, varav en del visas i figur 12, gör det rimligt att gissa att

$$0.001 \leq \frac{1}{\|K\|^2} \leq 0.01.$$

Låt oss se om vi kan göra en a priori uppskattning av operator- eller matrisnormen. *Frobeniusnormen* av en $m \times n$-matris A definieras som

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2},$$

en elementvis generalisering av den euklidiska normen på $\mathbb{R}^n$ till matrisrum.

Sats 3.2.1. *För en godtycklig $m \times n$-matris A gäller att*

$$\|A\| \leq \|A\|_F.$$

Bevis. Vi har att

$$\|A\|_F^2 = \sum_{i,j} a_{ij}^2 = \sum_i \|A_{i*}\|^2 = \sum_j \|A_{*j}\|^2,$$

där A_{i*} och A_{*j} betecknar den i:te raden respektive den j:te kolonnen i matrisen A. Cauchy-Schwarz olikhet ger därför att

$$\|Ax\|^2 = \sum_i (A_{i*}x)^2 \leq \sum_i \|A_{i*}\|^2\|x\|^2 = \|A\|_F^2\|x\|^2,$$

alltså att

$$\|Ax\| \leq \|A\|_F\|x\|,$$

för alla $x \in \mathbb{R}^n$. □

Skriver vi om x och k som kolonnvektorer får matrisen för den diskreta faltningsoperatorn K i två dimensioner samma form som i det endimensionella fallet, alltså kolonner med $k(:)$ i och för övrigt nollor. Matrisen är av typ $n \times m$, där

$$\begin{aligned} m &= \dim(x(:)) = m_1 m_2, \\ n &= \dim(y(:)) = (m_1 + l_1 - 1)(m_2 + l_2 - 1). \end{aligned}$$

Vi kan nu lätt beräkna Frobeniusnormen:

$$\|K\|_F^2 = m\|k(:)\|^2 \approx 570,$$

vilket ger oss att

$$\|K\|^2 \leq 570 < 10^3.$$

Denna uppskattning överensstämmer väl med våra experimentella resultat.

3.3 Tomografiproblemet och Radontransformen

Det inversa problem som presenteras i detta avsnitt är hämtat från medicinsk bildanalys, närmare bestämt från en matematisk-fysikalisk modell för Röntgentomografi. Denna modell motiveras med resonemang baserade på fysikalisk teori för hur Röntgenstrålning växelverkar med materien. Att närmare gå in på dessa bakomliggande resonemang eller detaljerna i modellen för oss dock utanför ämnet för den här uppsatsen. Det som ur matematisk synpunkt är av intresse för oss är den så kallade Radontransformen och det med denna förknippade inversa problem, vilket vi här kommer att benämna *tomografiproblemet*.

Ordet tomografi kommer från grekiskans *tomos* som betyder tvärsnitt. Så en översättning vore ungefär "bilder av tvärsnitt". Idag användes denna term för ett antal olika metoder för att rekonstruera den inre strukturen hos ett föremål utifrån yttre mätningar. I Röntgentomografin ligger vi nära ordets etymologiska rötter, emedan vi i praktiken gör oss en bild av ett tre-dimensionellt föremål med hjälp av dess två-dimensionella tvärsnitt.

Föremålet för den Röntgentomografiska analysen representeras i den matematiska modellen av en funktion $\mu : \mathbb{R}^3 \longrightarrow \mathbb{R}$, den så kallade *spridningskoefficienten*, vilket utgör ett mått på föremålets förmåga att i en given punkt sprida eller absorbera Röntgenstrålar av en given energi (en sorts täthetsfunktion alltså). Vi förutsätter att μ har kompakt stöd och är ickenegativ på hela $\mathbb{R}^3$. Anledningen till att μ är av intresse för medicinsk bildanalys är att olika anatomiska strukturer ofta uppvisar betydande olikheter i spridningskoefficienten. Ben har t.ex. mycket högre värden på μ än mjuk vävnad och olika sorters mjuk vävnad uppvisar också variationer i μ. För medicinska tillämpningar är det avgörande att sådana skillnader finns mellan frisk och cancersjuk vävnad.

Enligt den matematiska modellen ger oss de (idealiserade) Röntgentomografiska mätningarna eller Röntgenbilderna (de generaliserade) linjeintegralerna av spridningskoefficienten längs med olika räta linjer i $\mathbb{R}^3$. Tomografiapparaturen bestämmer ett koordinatsystem (x_1, x_2, x_3) för $\mathbb{R}^3$. Tvärsnitten till μ låter man konventionellt bestämmas genom att den sista koordinaten fixeras i koordinatsystemet. Vi antar att μ har ett kompakt stöd som är en delmängd av kuben $[-a, a]^3$, för något $a > 0$. För varje $c \in [-a, a]$ och varje rät linje ℓ i $\mathbb{R}^2$ ger mätningarna linjeintegralen av μ längs med linjen

$$\{(x_1, x_2, x_3) : x_3 = c \text{ och } (x_1, x_2) \in \ell\},$$

det vill säga

$$\int_\ell \mu(_, _, c) ds.$$

Med kännedom om alla linjeintegralerna försöker vi därefter rekonstruera tvärsnittet $x_3 = c$, eller med andra ord bestämma funktionen $f = \mu(_, _, c)$. De samlade två-dimensionella tvärsnitten för varierande c ger oss en uppfattning om det tre-dimensionella föremålet.

Det inversa problem som modellen leder fram till är alltså följande: låter sig en funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ rekonstrueras utifrån kännedom om linjeintegralerna av f längs med alla räta linjer i planet? Man kan visa att detta i princip låter sig göras för alla i detta sammanhang relevanta funktioner. I praktiken har vi förstås endast kännedom om linjeintegralerna för ett ändligt antal linjer och begränsad precision i mätningarna. Det vi dock kan hoppas på är att åtminstone kunna göra en god approximation.

Mängden av räta linjer i planet $\mathbb{R}^2$ kan parametriseras med hjälp av en enhetsvektor $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$, där $\omega_1^2 + \omega_2^2 = 1$, och ett $t \in \mathbb{R}$. Vi definierar linjen

$$\ell_{t,\omega} := \{(x_1, x_2) \in \mathbb{R}^2 : \langle (x_1, x_2), \omega \rangle = t\}.$$

Vektorn ω kallas för den *ortogonala riktningen* för linjen $\ell_{t,\omega}$ och talet t linjens *affina parameter*. Linjer med samma ortogonala riktning är parallella och vinkelräta mot ω. Absolutbeloppet av den affina parametern $|t|$ är avståndet från linjen till koordinatsystemets origo. Figur 13 ger en illustration av begreppen.

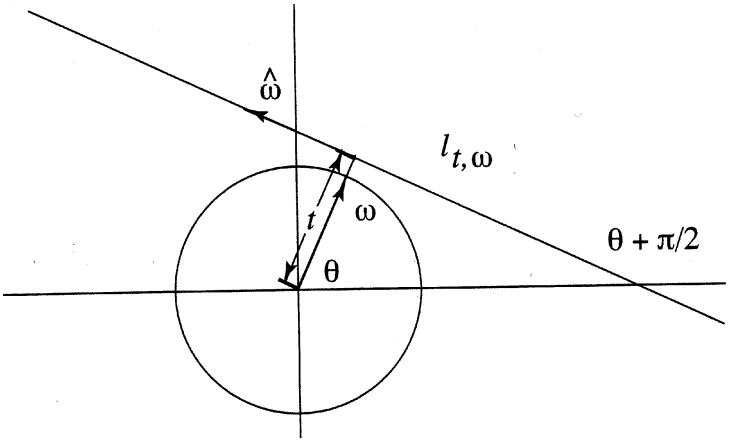
Som mängder betraktade är

$$\ell_{t,\omega} = \ell_{-t,-\omega},$$

eftersom $\langle (x, y), -\omega \rangle = -t$ omm $\langle (x, y), \omega \rangle = t$. Det ordnade paret (t, ω) bestämmer dock entydigt en *riktad* rät linje i planet enligt en orienteringskonvention. Sätt

$$\hat{\omega} = (-\omega_2, \omega_1).$$

Denna enhetsvektor samt $-\hat{\omega}$ är vinkelräta mot ω och därmed parallella med $\ell_{t,\omega}$. Vektorn $\hat{\omega}$ bestämmer linjens *positiva* riktning och $-\hat{\omega}$ dess *negativa*



Figur 13: Parametrisering av rummet av riktade räta linjer i planet. Bilden är hämtad ur [1].

riktning. Det omvända förhållandet gäller för linjen $\ell_{-t,-\omega}$. Det råder med andra ord ett-till-ett-korrespondens mellan mängden $\mathbb{R} \times S^1$ och rummet Λ av orienterade räta linjer i planet, där S^1 är enhetscirkeln

$$S^1 = \{\omega = (\omega_1, \omega_2) \in \mathbb{R}^2 : \omega_1^2 + \omega_2^2 = 1\}.$$

Alltså, om (t, ω) och (t', ω') är distinkta punkter i $\mathbb{R} \times S^1$ och $\ell_{t,\omega} = \ell_{t',\omega'}$, så är $(t', \omega') = (-t, -\omega)$.

Linjen $\ell_{t,\omega}$ kan parameterframställas som punktmängden

$$\ell_{t,\omega} = \{t\omega + s\hat{\omega} : s \in \mathbb{R}\}.$$

Vi kan också parametrisera enhetscirkeln S^1 med ett $\theta \in \mathbb{R}$, nämligen vinkeln för vektorn ω:

$$\omega(\theta) = (\cos\theta, \sin\theta).$$

Vi använder även beteckningen $\ell_{t,\theta}$ för linjen $\ell_{t,\omega}$.

Definition 3.3.1. *Låt $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ vara kontinuerlig med kompakt stöd.* Radontransformen av f *är avbildningen $\mathcal{R}f : \mathbb{R} \times S^1 \longrightarrow \mathbb{R}$, där*

$$\mathcal{R}f(t, \omega) = \int_{\ell_{t,\omega}} f ds.$$

Radontransformen *är den avbildning $\mathcal{R}$ mellan funktionsrum som är sådan att*

$$\mathcal{R}(f) = \mathcal{R}f,$$

och vars definitionsrum är det linjära rummet

$$D_{\mathcal{R}} = \{f : \mathbb{R}^2 \longrightarrow \mathbb{R} : f \text{ har kompakt stöd}\}.$$

Observera att

$$\int_{\ell_{t,\omega}} f ds = \int_{-\infty}^{\infty} f(s\hat{\omega} + t\omega) ds,$$

vilket vi ser genom att utnyttja parameteriseringen

$$s \in \mathbb{R} \mapsto \gamma(s) = s\hat{\omega} + t\omega$$

och beräkna linjeintegralen. Vi får att

$$\int_{\ell_{t,\omega}} f ds = \lim_{\tau \to \infty} \int_{-\tau}^{\tau} f(\gamma(s))|\gamma'(s)| ds = \int_{-\infty}^{\infty} f(\gamma(s)) ds,$$

då

$$|\gamma'(s)| = \sqrt{\left(\frac{dx_1}{ds}\right)^2 + \left(\frac{dx_2}{ds}\right)^2} = \sqrt{(-\omega_2)^2 + \omega_1^2} = 1.$$

Att Radontransformen av f är väldefinierad på hela $\mathbb{R} \times S^1$ följer av att f antas ha kompakt stöd. I medicinska tillämpningar uppfyller (vanligen) de relevanta funktionerna det senare villkoret samt är styckvis kontinuerliga, varför utnyttjande av Radontransformen är möjligt.

Sats 3.3.1.

(i) Radontransformen är en linjär avbildning (mellan funktionsrum):

$$\mathcal{R}(af + bg) = a\mathcal{R}f + b\mathcal{R}g,$$

för alla $a,\ b \in \mathbb{R}$ *och* $f,\ g \in D_{\mathcal{R}}$.

(ii) Radontransformen är monoton: om f *är en icke-negativ funktion i* $D_{\mathcal{R}}$*, så är*

$$\mathcal{R}f(t, \omega) \geq 0.$$

(iii) Radontransformen av f *är en jämn funktion:*

$$\mathcal{R}f(t, \omega) = \mathcal{R}f(-t, -\omega).$$

Bevis. Fall (i) följer av linjeintegralens linjäritet och (ii) återigen från linjeintegralens egenskaper. Fall (iii) följer av att som mängder betraktade, alltså bortsett från orienteringen, är $\ell_{t,\omega} = \ell_{-t,-\omega}$. □

Vi är nu redo att ge den fullständiga matematiska formuleringen av det inversa problemet.

> **Tomografiproblemet:** Givet Radontransformen $\mathcal{R}f$ av en kontinuerlig icke-negativ funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ med kompakt stöd, består problemet i att bestämma f.

Man kan visa att detta låter sig göras för funktioner av nämnd typ. Helt generellt ges då f av ett integraluttryck i $\mathcal{R}f$. Det finns med andra ord en linjär avbildning $\mathcal{R}^{-1}$ från reellvärda funktioner på $\mathbb{R} \times S^1$ till reellvärda funktioner på $\mathbb{R}^2$ som uppfyller

$$\mathcal{R}^{-1} \circ \mathcal{R}(f) = f.$$

Beviset för detta är dock något invecklat och involverar Fouriertransformen, så vi utelämnar det (för mer information, se [1]).

Låt oss nu ta itu med diskretiseringen av tomografiproblemet. Liksom i föregående avsnitt diskretiseras funktionen f över någon "stödrektangel" $D \subset \mathbb{R}^2$, som kan indelas i $m \times n$ kvadrater med sidan Δx. För varje sådan kvadrat låter vi x_{ij}, där $1 \leq i \leq m$ och $1 \leq j \leq n$, vara mittpunkten på kvadraten. Detta ger en diskretisering av D i $m \times n$ jämnt fördelade punkter med avståndet Δx emellan, både vertikalt och horisontellt. Dessa punkter sammanställes i en matris

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}.$$

Funktionen f representeras av matrisen

$$f = \begin{bmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & & \vdots \\ f_{m1} & \cdots & f_{mn} \end{bmatrix},$$

där $f_{ij} = f(x_{ij})$ för alla par (i, j). Matrisen f är så att säga $''f(X)''$. Låt oss för enkelhetens skull förutsätta att m och n är udda tal. Vi kan då göra antagandet att matrisen av punkter X har sitt centrum x_{uv} i origo, där

$$u = \frac{m+1}{2} \quad \text{och}$$
$$v = \frac{n+1}{2}.$$

Annars kan vi alltid göra ett koordinatbyte. Detta ger

$$x_{ij} = ((i-u)\Delta x, (j-v)\Delta x),$$

för $1 \leq i \leq m$ och $1 \leq j \leq n$.

Radontransfomen av f representeras som en matris Rf av typ $M \times N$, där M är antalet element i vektorn

$$\bar{\theta} = (0, \frac{\pi}{M}, \frac{2\pi}{M}, \ldots, \frac{(M-1)\pi}{M}),$$

en diskretisering av intervallet $[0, \pi)$ i ett lämpligt antal jämnt fördelade punkter (eftersom Radontransformen av f är en jämn funktion behöver vi inte titta på hela vinkelintervallet $[-\pi, \pi]$). Talet N är antalet element i vektorn $\bar{t} = (t_1, \ldots, t_N)$, där

$$t_J = t_1 + (J-1)\Delta x,$$

för $1 \leq J \leq N$. Denna vektor utgör en diskretisering av ett ändligt intervall $[a, b]$ i $\mathbb{R}$ som är sådant att varje linje som passerar genom rektangeln D kan framställas på formen $\ell_{t,\theta}$ för något $t \in [a, b]$ och någon vinkel $\theta \in [0, \pi)$.

Vi behöver göra ett lämpligt val av a och b. Låt $\theta \in [0, \pi)$ vara godtyckligt och låt PQ vara det segment av linjen $\{s\omega(\theta) : s \in \mathbb{R}\}$ som utgöres av den ortogonala projektionen av rektangeln D på nämnda linje. En övre gräns för $\|PQ\|$ utgöres av diametern på mängden D, som för en rektangel är lika med längden på diagonalen:

$$\|PQ\| \leq \operatorname{diam}(D) = \sqrt{(m^2 + n^2)}\Delta x,$$

Eftersom rektangeln är centrerad i origo behöver vi inte sätta b till hela diagonalen, utan kan välja en bråkdel av den, lämpligen något större än halva diagonalen för att få lite marginal. Vi sätter alltså

$$b = w \cdot \operatorname{diam}(D), \quad \text{där } w > 0.5,$$

och $a = -b$.

Ett lämpligt val av N är

$$N = \left\lceil \frac{b-a}{\Delta x} \right\rceil + 1 = \left\lceil \frac{2w \cdot \operatorname{diam}(D)}{\Delta x} \right\rceil + 1,$$

där $\lceil _ \rceil$ är den på $\mathbb{R}$ definierade *takfunktionen*

$$\lceil c \rceil = \min_{p \in \mathbb{Z}} \{p \geq c\}.$$

Med valet $t_1 = a$ får vi att

$$t_{N-1} \leq b \leq t_N,$$

emedan $\lceil c \rceil - 1 \leq c \leq \lceil c \rceil$, för alla $c \in \mathbb{R}$.

Låt nu $Rf_{\theta t}$ beteckna ett godtyckligt element i matrisen Rf. Vi vill bestämma detta element så att

$$\int_{\ell_{t,\theta}} f ds \approx Rf_{\theta t}.$$

Till varje linje $\ell_{t,\theta}$ kan vi förknippa en matris $\Delta S = \Delta S(\theta, t)$ av typ $m \times n$, sådan att: om linjen $\ell_{t,\theta}$ passerar en kvadrat med sidan Δx och med mittpunkten x_{ij}, så utgör elementet $\Delta s_{ij} = \Delta s_{ij}(\theta, t)$ i ΔS en approximation av

längden på det segment av $\ell_{t,\theta}$ som passerar genom nämnda kvadrat; annars är $\Delta s_{ij} = 0$. Vi sätter

$$Rf_{\theta t} := \sum_{i,j} f_{ij} \Delta s_{ij} = \Delta S(:)^T f(:).$$

Denna formel gör det lätt att konstruera en matris $[R]$ av typ $(MN) \times (mn)$, sådan att

$$Rf(:) = [R] f(:).$$

Raderna i $[R]$ utgöres av vektorerna $\Delta S(\theta, t)(:)^T$ för varierande θ och t. Vi har här följaktligen en linjär avbildning

$$R : \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}^{M \times N},$$

som får tjäna som diskret representation av Radontransformen.

Utan vidare motivering konstaterar vi bara att en lämplig approximation av längden på linjesegmentet är

$$\Delta s_{ij} = \begin{cases} \Delta x - \mathrm{d}(x_{ij}, \ell_{t,\theta}) & \text{om denna differens är positiv,} \\ 0 & \text{annars.} \end{cases}$$

Den andra termen i differensen är avståndet från mittpunkten x_{ij} till linjen. Låt $z = (z_1, z_2)$ vara den punkt på linjen som är sådan att

$$\mathrm{d}(x_{ij}, \ell_{t,\theta}) = \|z - x_{ij}\|.$$

Vi har att

$$\begin{aligned} \|z - x_{ij}\| &= |\langle z - x_{ij}, \omega(\theta) \rangle| = |\langle z, \omega(\theta) \rangle + \langle x_{ij}, \omega(\theta) \rangle| \\ &= |t - \langle x_{ij}, \omega(\theta) \rangle|, \end{aligned}$$

vilket sammanfattningsvis ger

$$\mathrm{d}(x_{ij}, \ell_{t,\theta}) = |t - \langle x_{ij}, \omega(\theta) \rangle|.$$

I beräkningen av matrisen $[R]$ kommer vi att gå igenom alla $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$ och $I \in \{1, \ldots, M\}$, svarande mot θ_I. För givna i, j och I behöver vi dock inte gå igenom alla t-värden, eftersom flertalet av dessa ger $\Delta s_{ij}(\theta, t) = 0$. Vi har att $\Delta s_{ij} > 0$ omm

$$\begin{aligned} &\Delta x - |t - \langle x_{ij}, \omega(\theta) \rangle| > 0 \quad \Longleftrightarrow \\ &-\Delta x < t - \langle x_{ij}, \omega(\theta) \rangle < \Delta x \quad \Longleftrightarrow \\ &\langle x_{ij}, \omega(\theta) \rangle - \Delta x < a + (J-1)\Delta x < \langle x_{ij}, \omega(\theta) \rangle + \Delta x, \end{aligned}$$

där $1 \leq J \leq N$. Sätter vi

$$c = \frac{1}{\Delta x}(\langle x_{ij}, \omega(\theta)\rangle - a),$$

får vi att

$$c < J < c + 2.$$

Eftersom J är ett heltal ger detta antingen en eller två möjliga lösningar, en lösning blott i det fallet c är ett heltal. Vi kan ignorera detta och tänka att istället för att lösa den stränga olikheten, tittar vi på olikheten $c < J \leq c+2$, vilket för alla fall ger de två heltalslösningarna

$$J = [c] + 1,\ [c] + 2,$$

där $[_] : \mathbb{R} \longrightarrow \mathbb{Z}$ är heltalsdelen

$$[c] = \max_{p \in \mathbb{Z}} \{p \leq c\}.$$

Vi ser att positionen (i, j) i en $m \times n$-matris A korresponderar mot positionen $m(j-1)+i$ i kolonnvektorn $A(:)$. Detta betyder för vår matris $[R]$ att paret (i, j) korresponderar mot kolonn nummer $m(j-1)+i$ i $[R]$ och att paret (I, J) korresponderar mot rad nummer $M(J-1)+I$.

Matrisen [R] är vad som i numerisk analys kallas för en *gles* matris, vilket innebär att den huvudsakligen består av nollor. Det finns algoritmer för att på ett kompakt sätt representera en sådan matris som en särskild datastruktur, där grundtanken är att man enbart lagrar de nollskilda elementen och deras positioner i matrisen. I *MatLab* kan vi göra detta med kommandot

$$S = \text{sparse}(m, n),$$

varvid utrymme allokeras för en gles matris S av typen $m \times n$. Därefter kan elementen i S anges på samma sätt som i det vanliga fallet. Elementen sättes till noll om inget annat anges (nollmatrisen kallas i detta sammanhang också för den "ultimata" glesa matrisen). Den till S transponerade matrisen är också gles. Däremot behöver produkten av två glesa matriser inte vara särskilt gles och inte heller inversen. Därför vill vi i beräkningar med S undvika att t.ex. beräkna matrisinversen; för detta utnyttjar vi liksom i föregående avsnitt istället *MatLab*-funktionen pcg, som också accepterar datastrukturen gles matris. Och en produkt på formen STx, där x är en kolonnmatris, beräknas i ordningen $S(Tx)$ och inte i ordningen $(ST)x$ (det gäller att sätta parenteserna rätt!).

Pseudokod för konstruktion av matrisen $[R]$ är som följer:

function $R = \text{radon}(m, n, \Delta x, M)$
  global N
    $\gamma = \text{diam}(D)$
    $b = w\gamma;\ a = -b$
    $N = \lceil 2w\gamma \rceil + 1$
    $u = (m+1)/2;\ v = (n+1)/2$
    $R = \text{sparse}(MN, mn)$
  for $i = 1$ to m, $\quad j = 1$ to n, $\quad I = 1$ to M
    $x = ((i-u)\Delta x, (j-v)\Delta x)$
    $\theta = (I-1)\pi/M; \quad \omega = (\cos\theta, \sin\theta)$
    $d = \langle x, \omega \rangle; \quad c = [(b+d)/\Delta x] + 1$
  for $J = c : c+1$
    $t = a + (J-1)\Delta x$
    $k = M(J-1) + I; \quad l = m(j-1) + i$
    $R(k,l) = \Delta x - |t-d|$
  end
  end.

Denna algoritm, eller snarare detta sätt att formulera algoritmen, blir dock väldigt långsam i *MatLab*. Vi får en avsevärt mycket snabbare algoritm ifall vi effektivt utnyttjar *MatLabs* inbyggda förmåga att hantera matriser/vektorer. För $m = n = 101$ och $M = 100$ tog det för ovanstående algoritm ungefär en timme att beräkna $[R]$, medan det för den snabbare gick på cirka 45 sekunder. Nedan ges *MatLab*-pseudokoden för den snabbare algoritmen:

function $R = \text{snabbRadon}(m, n, \Delta x, M)$
  global N
    $\gamma = \text{diam}(D)$
    $b = w\gamma;\ a = -b$
    $N = \lceil 2w\gamma \rceil + 1$
    $u = (m+1)/2;\ v = (n+1)/2$
    $R = \text{sparse}(MN, mn)$
    $x = ((1:m) - u)\Delta x;\ y = ((1:n) - v)\Delta x$
    $[X, Y] = \text{ndgrid}(x, y)$
  for $I = 1$ to M
    $\theta = (I-1)\pi/M; \quad \omega = (\cos\theta, \sin\theta)$
    $d = \omega_1 X + \omega_2 Y; \quad c = [(b+d)/\Delta x] + 1$
    $t = a + (c-1)\Delta x$
    $k = (c-1)M + I;\ l = 1:mn$
    $R = R + \text{sparse}(k(:), l(:), \Delta x - |t(:) - d(:)|, MN, mn)$
    $t = a + c\Delta x$
    $k = cM + I;\ l = 1:mn$
    $R = R + \text{sparse}(k(:), l(:), \Delta x - |t(:) - d(:)|, MN, mn)$
  end.

3.3.1 Numeriska experiment med tomografiproblemet

För alla de tre regulariseringsmetoderna kan vi med små justeringar återanvända *MatLab*-kod från de föregående två avsnitten. Som exakt lösning använder vi en så kallad *Shepp-Logan-fantom*, vilket är en vanlig matematisk modell för de tvärsnitt av människokroppen man stöter på i datortomografin (se [1]). *MatLab*-kod för att generera en lämplig fantom finns att hämta på sidan [5]. Vi sätter $m = n = 101$, $M = 100$ och $\Delta x = 1$. Vår Shepp-Logan-matris har reellvärda element i intervallet $[0, 2]$ och ritas i *MatLab* enligt en gråskala med hjälp av kommandona

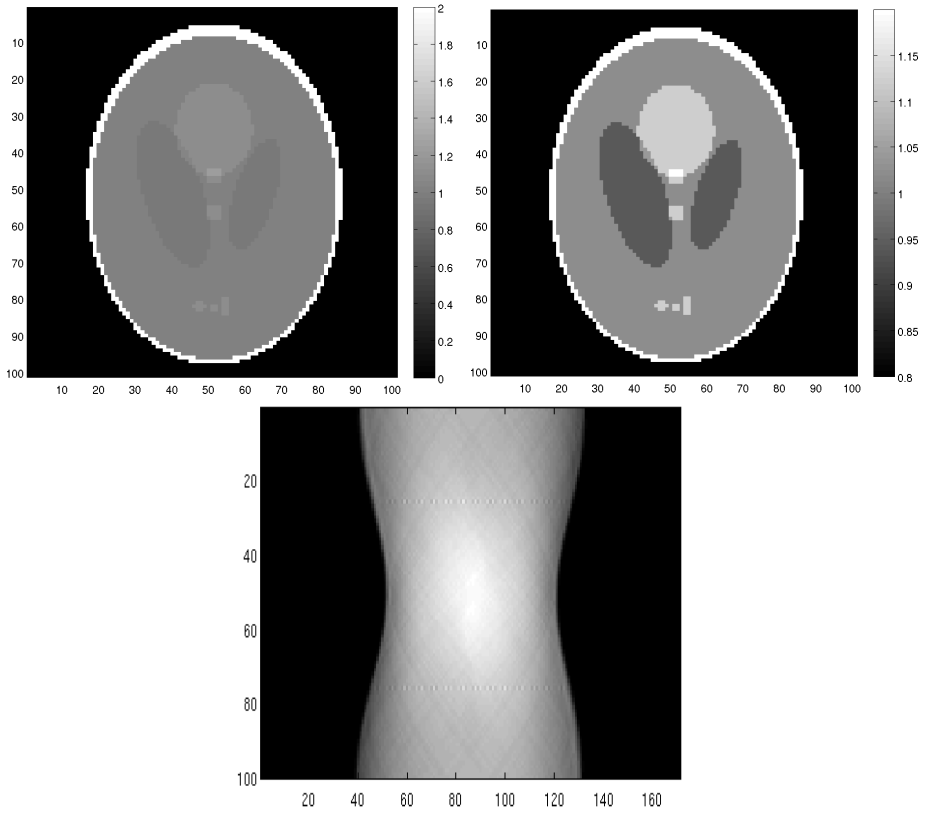
$$\text{imagesc}(f);\ \text{axis image};\ \text{colormap(gray)};\ \text{colorbar}.$$

För att få mer kontrast har vi parallellt med originalet också ritat alla ”fantombilder” med kommandot

$$\text{imagesc}(f, [0.8, 1.2]),$$

vilket gör så att alla värden i matrisen under 0.8 reduceras till svart ($= 0$) och alla värden över 1.2 till vitt ($= 2$). Värdeintervallet visas i färgstapeln

till höger om bilderna. I figur 14 visas den exakta lösningen f och det perturberade högerledet y^δ, det vill säga den diskretiserade Radontransformen av f plus någon störning såsom i ekvation (36). Radontransformen har värden i intervallet $[0, 100]$ och vi har satt $\epsilon = 0.5$.



Figur 14: Övre bilder: Shepp-Logan-fantom ritad med mindre och mer kontrast. Undre bild: Radontransformen av fantomen, med vinkeln θ på y-axeln och den affina parametern på x-axeln.

Tikhonovregularisering. Denna utföres med hjälp av funktionen pcg på samma sätt som i avsnitt 3.2.1. Pseudokod för motsvarande funktioner blir i detta fall:

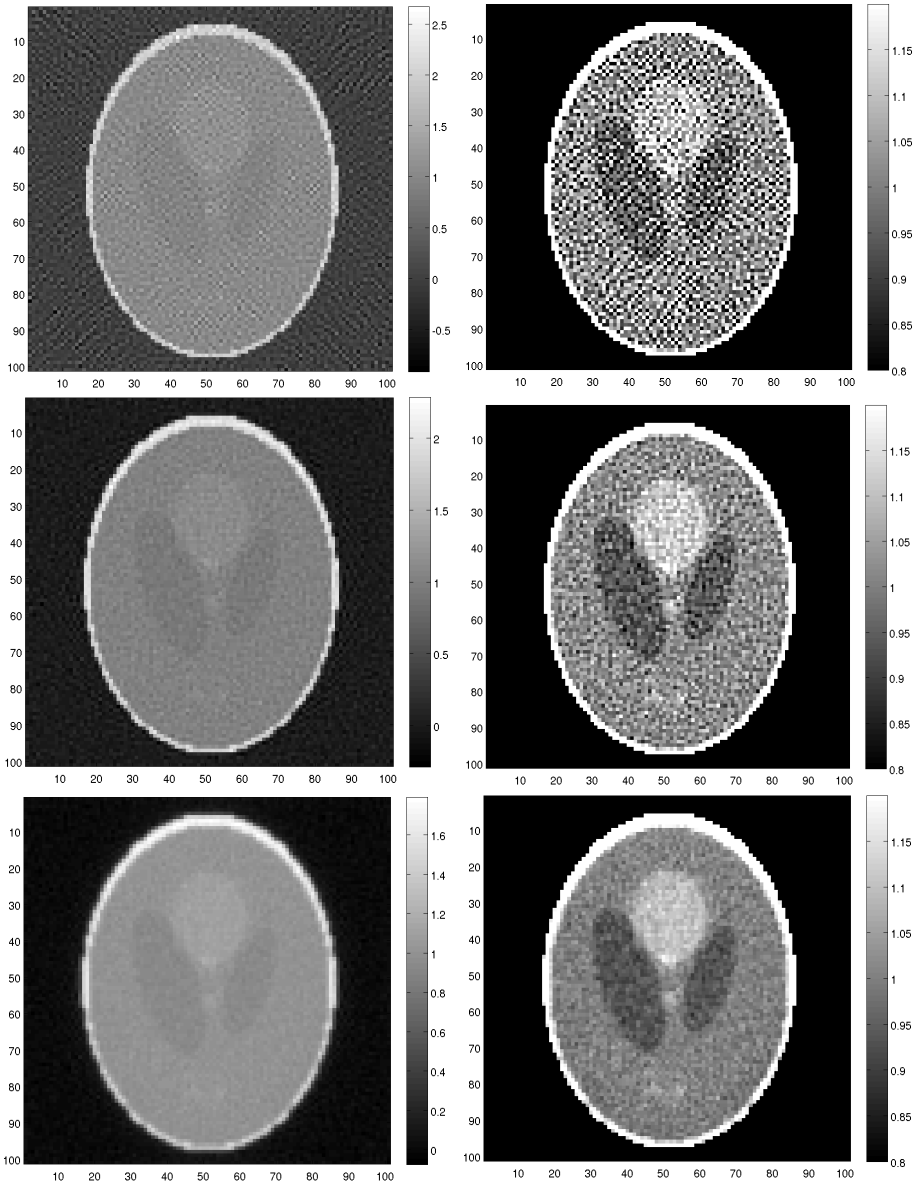
$$\begin{aligned} &\text{function } y = F(x) \\ &\quad \text{global } R,\ \alpha' \\ &\qquad y = \alpha' x + R^t(Rx) \end{aligned}$$

function $x^\alpha = T(\alpha)$
  global $R,\ \alpha'$
    $\alpha' = \alpha$
    $z = R^t y$
    $x = \text{pcg}(F, z)$
    $x = \text{reshape}(x, m, n).$

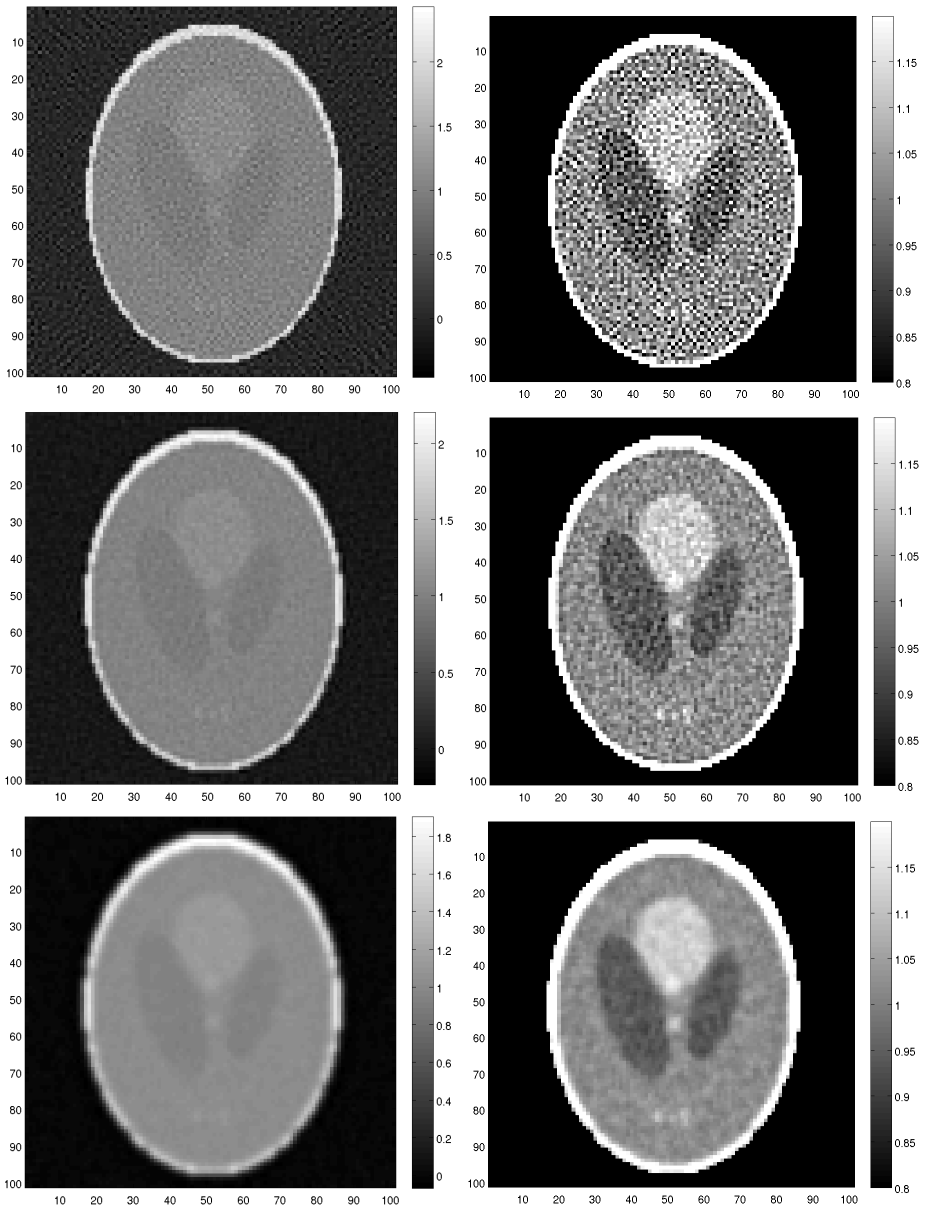
Generaliserad Tikhonovregularisering. Här återanvänder vi *MatLab*-koden för funktionen L_2 från avsnitt 3.2.1 och modifierar ovanstående kod för funktionen F.

Landweberiteration. Frobeniusnormen av $[R] \approx 825$, vilket betyder att ett val av parametern a med $0 < a \leq 10^{-6}$ gör iterationen möjlig. Erfarenheten tycks dock visa att iterationen fungerar bra med $a = 0.00005$ också. Pseudokoden för iterationen är rakt på sak:

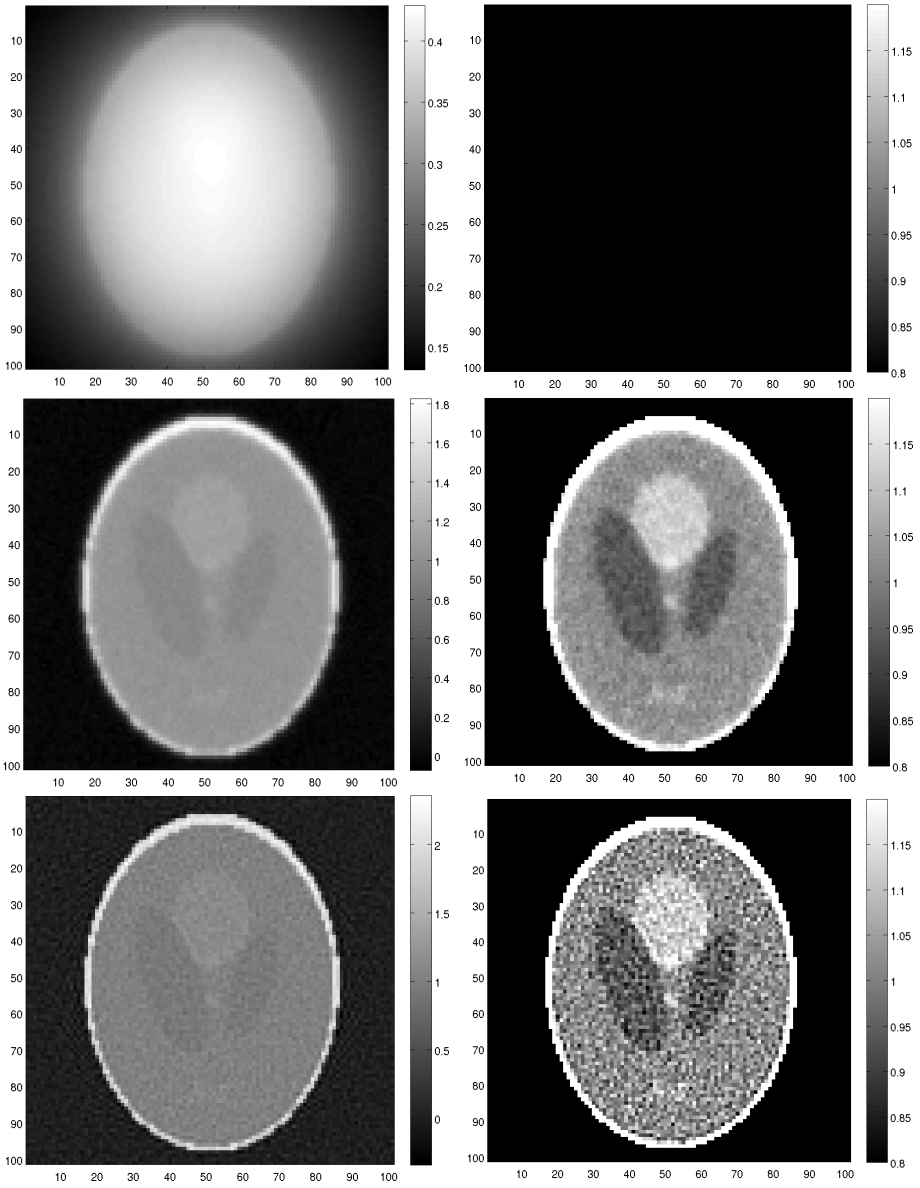
function $x = \text{Lw}(p, a)$
    $x = \text{zeros}(mn, 1)$
  for $i = 1$ to p
    $x = x - aR^t(Rx) + aR^t y$
  end
    $x = \text{reshape}(x, m, n).$



Figur 15: Tikhonovregularisering. I kolumnen till höger ritas resultatet med mer kontrast. Uppifrån och ned: $\alpha = 0.1$, $\alpha = 10$ och $\alpha = 100$.



Figur 16: Generaliserad Tikhonovregularisering. I kolumnen till höger ritas resultatet med mer kontrast. Uppifrån och ned: $\alpha = 0.1$, $\alpha = 10$ och $\alpha = 100$.



Figur 17: Landweberiteration med parametern $a = 0.00005$. I kolumnen till höger ritas resultatet med mer kontrast. Uppifrån och ned: antal iterationer $p = 1$, $p = 100$ och $p = 1000$.

Referenser

[1] C. L. Epstein. *Introduction to the Mathematics of Medical Imaging.* Pearson Education/Prentice Hall, New Jersey, 2003.

[2] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems.* Applied Mathematical Sciences, volume 120, Springer-Verlag, New York, 1996.

[3] S. Lang. *Linear Algebra.* Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1987.

[4] C. R. Vogel. *Computational Methods for Inverse Problems.* SIAM 2002.

[5] http: //server.oersted.dtu.dk/ftp/jaj/31655/ct_programs/