Laboration 2: Icke-linjär parameterestimering med bivillkor

Niclas Börlin

8 december 1999

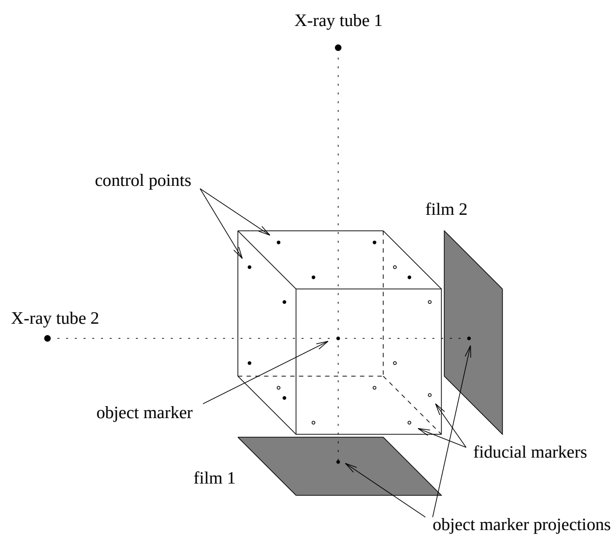
1 Inledning

I denna laboration ska ni implementera en metod för sekvensiell kvadratisk programmering och applicera den på olika bivillkorsformuleringar av ett projektionsproblem.

2 Bakgrund

Röntgenstereofotogrammetri (*Radiostereometric Analysis — RSA*) är en metod att mäta mikrorörelser i skelettet.

Vid en RSA-undersökning exponeras patienten med två röntgenrör samtidigt tillsammans med en referensbur. Inopererade markörer av tantal (0.5–1.0 mm∅) avbildas på bägge röntgenfilmerna. Markörer i referensburen avbildas på respektive film. (Figur 1.)



Figur 1: Vid en RSA-undersökning exponeras patienten och en referensbur med två röntgenrör samtidigt. Patientmarkörer avbildas på bägge filmerna medan referensmarkörer i buren avbildas på resp. film.

Ett av delproblemen i RSA är att rekonstruera 3d-koordinaterna för patientkulorna. Då måste man först rekonstruera projektionsgeometrin för de två fokusen.

2.1 Modell

Geometrin motsvarar en *centralprojektion* då alla strålar går genom en centralpunkt eller fokus.

Algebraiskt avbildas en objektpunkt $p = [x, y, z]^T$ på en filmpunkt $q = [u, v]^T$ enligt

$$\begin{bmatrix} u \\ v \\ 0 \end{bmatrix} - \begin{bmatrix} u_0 \\ v_0 \\ c \end{bmatrix} = R \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \right). \tag{1}$$

Punkten $f = [f_x, f_y, f_z]^T$ är fokuskoordinaterna i världskoordinatsystemet. Matrisen R en rotationsmatris mellan världs- och bildkoordinatsystemet. Punkten $[u_0, v_0]^T$ är origo (principalpunkt) i bildkoordinatsystemet och konstanten c är film-fokus-avståndet (fokalavstånd, principalavstånd). Tillsammans beskriver $[u_0, v_0, c]^T$ fokuskoordinaterna i bildkoordinatsystemet.

Formuleringen (1) kan kallas den *geometriska* formuleringen av projektionen, då fokus, etc., ingår explicit i projektionsformlerna. Den har 9 frihetsgrader; fokus i de två koordinatsystemen $[f_x, f_y, f_z]^T$ och $[u_0, v_0, c]^T$ samt 3 rotationsvinklar mellan koordinatsystemen.

Om vi löser ut och sätter in c i (1) får vi

$$\begin{bmatrix} u \\ v \end{bmatrix} - \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \frac{-c}{r_{31}(x - f_x) + r_{32}(y - f_y) + r_{33}(z - f_z)} \begin{bmatrix} r_{11}(x - f_x) + r_{12}(y - f_y) + r_{13}(z - f_z) \\ r_{21}(x - f_x) + r_{22}(y - f_y) + r_{23}(z - f_z) \end{bmatrix}$$

eller

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} - \frac{\begin{bmatrix} c(r_{11}(x - f_x) + r_{12}(y - f_y) + r_{13}(z - f_z)) \\ c(r_{21}(x - f_x) + r_{22}(y - f_y) + r_{23}(z - f_z)) \end{bmatrix}}{r_{31}(x - f_x) + r_{32}(y - f_y) + r_{33}(z - f_z)}$$

vilket kan skrivas om till

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{\begin{bmatrix} a_{11}x + a_{12}y + a_{13}z + a_{14} \\ a_{21}x + a_{22}y + a_{23}z + a_{24} \end{bmatrix}}{a_{31}x + a_{32}y + a_{33}z + a_{34}} = \frac{\begin{bmatrix} A_1^T p + a_{14} \\ A_2^T p + a_{24} \end{bmatrix}}{A_3^T p + a_{34}}$$

eller

$$q' = \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = A \begin{bmatrix} p \\ 1 \end{bmatrix}, \; q = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u'/w' \\ v'/w' \end{bmatrix} \tag{2}$$

där a_{ij} är olika konstanter som ej beror av x, y, z:

$$A = \begin{bmatrix} -cr_{11} + u_0 r_{31} & -cr_{12} + u_0 r_{32} & -cr_{13} + u_0 r_{33} & cr_{11} f_x + cr_{12} f_y + cr_{13} f_z - u_0 r_{31} f_x - u_0 r_{32} f_y - u_0 r_{33} f_z \\ -cr_{21} + v_0 r_{31} & -cr_{22} + v_0 r_{32} & -cr_{23} + v_0 r_{33} & cr_{21} f_x + cr_{22} f_y + cr_{23} f_z - v_0 r_{31} f_x - v_0 r_{32} f_y - v_0 r_{33} f_z \\ r_{31} & r_{32} & r_{33} & -r_{31} f_x - r_{32} f_y - r_{33} f_z \end{bmatrix}.$$

Formulering (2) kan kallas den *algebraiska* formuleringen då de geometriska parametrarna ej ingår explicit. Den algebraiska formuleringen har 12 frihetsgrader, 3 fler än den geometriska. Dessa 3 extra frihetsgrader går att tolka som att vi tillåter

1. en global skalfaktor, dvs om vi byter enhet i ett koordinatsystem så byter vi också enhet i det andra,
2. ett "skalfel" mellan de två axlarna i bildkoordinatsystemet och
3. icke-ortogonalitet mellan de två axlarna i bildkoordinatsystemet.

2.2 Det inversa problemet

Vårt problem är att givet m kända 3d-koordinater p_i och dess uppmätta 2d-koordinater q_i bestämma parametrarna för projektionen, dvs lösa

$$\begin{array}{rl} \min\limits_{A} & \frac{1}{2}\sum\limits_{i=1}^{m} \|q_i - \tilde{q}_i\|^2 \\ \text{sådan att} & \text{A centralprojektion } (q_i \text{ bestäms enligt (2)}) \end{array} \tag{3}$$

och ev. ytterligare bivillkor.

2.3 Global skalfaktor

Den globala skalfaktorn är omöjlig att bestämma från enbart kända och uppmätta koordinater och är dessutom ointressant. Den måste därför elimineras med hjälp av en normering av A. Vi kommer att använda två olika normeringsbivillkor

$$a_{34} = 1 \tag{4}$$

$$A_3^T A_3 = 1 \tag{5}$$

2.4 Skalfel

Att vi saknar skalfel i mätkoordinatsystemet motsvarar bivillkoret

$$(A_1^T A_1 - A_2^T A_2)A_3^T A_3 - \left((A_1^T A_3)^2 - (A_2^T A_3)^2\right) = 0 \tag{6}$$

2.5 Ortogonalitet mellan mätaxlarna

Att vi har ortogonalitet mellan mätaxlarna motsvarar bivillkoret

$$(A_1^T A_2)(A_3^T A_3) - (A_1^T A_3)(A_2^T A_3) = 0 \tag{7}$$

3 Startapproximationer

För detta problem finns det ett relativt enkelt sätt att bestämma en bra startpunkt. Om vi antar att vi saknar mätfel, dvs $\tilde{q}_i = q_i$, så gäller att

$$\begin{aligned}
\tilde{q}_i = q_i &= \frac{1}{A_3^T p_i + a_{34}} \begin{bmatrix} A_1^T p_i + a_{14} \\ A_2^T p_i + a_{24} \end{bmatrix} \\
\tilde{q}_i (A_3^T p_i + a_{34}) &= \begin{bmatrix} A_1^T p_i + a_{14} \\ A_2^T p_i + a_{24} \end{bmatrix} \\
\begin{bmatrix} \tilde{q}_{i1} A_3^T p_i + \tilde{q}_{i1} a_{34} \\ \tilde{q}_{i2} A_3^T p_i + \tilde{q}_{i2} a_{34} \end{bmatrix} &= \begin{bmatrix} A_1^T p_i + a_{14} \\ A_2^T p_i + a_{24} \end{bmatrix}.
\end{aligned}$$

Om vi använder normaliseringsvillkor (4) får vi

$$\begin{aligned}
\begin{bmatrix} \tilde{q}_{i1} A_3^T p_i + \tilde{q}_{i1} \\ \tilde{q}_{i2} A_3^T p_i + \tilde{q}_{i2} \end{bmatrix} &= \begin{bmatrix} A_1^T p_i + a_{14} \\ A_2^T p_i + a_{24} \end{bmatrix} \\
\begin{bmatrix} -p_i^T A_1 - a_{14} + \tilde{q}_{i1} p_i^T A_3 \\ -p_i^T A_2 - a_{24} + \tilde{q}_{i2} p_i^T A_3 \end{bmatrix} &= \begin{bmatrix} -\tilde{q}_{i1} \\ -\tilde{q}_{i2} \end{bmatrix} = -\tilde{q}_i \\
&\text{eller} \\
\underbrace{\begin{bmatrix} -p_i^T & -1 & 0 & 0 & \tilde{q}_{i1} p_i^T \\ 0 & 0 & -p_i^T & -1 & \tilde{q}_{i2} p_i^T \end{bmatrix}}_{B_i} \underbrace{\begin{bmatrix} A_1 \\ a_{14} \\ A_2 \\ a_{24} \\ A_3 \end{bmatrix}}_{x_0'} &= -\tilde{q}_i.
\end{aligned}$$

Med m punkter $(p_i, \tilde{q}_i)$ kan vi då lösa

$$\min_{x_0'} \quad \|Bx_0' - b\|^2, \tag{8}$$

$$\text{där} \quad B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix}, \; b = \begin{bmatrix} -\tilde{q}_1 \\ -\tilde{q}_2 \\ \vdots \\ -\tilde{q}_m \end{bmatrix}$$

och sedan använda $x_0 = \begin{bmatrix} x_0' \\ 1 \end{bmatrix}$ som startapproximation.

Har vi något annat normeringsvillkor kan vi först lösa (8) och sedan skala om x_0.

4 Uppgifter

4.1 Grunduppgift

- Implementera sekvensiell kvadratisk programmering (SQP) där ni löser systemekvationen

$$\begin{bmatrix} 0 & 0 & A \\ 0 & I & J \\ A^T & J^T & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ r \\ p \end{bmatrix} = \begin{bmatrix} -g \\ -F \\ 0 \end{bmatrix} \tag{9}$$

  och tar odämpade steg. Använd $\|Jp\| \leq \epsilon_F(1 + \|F\|)$ och $|g| \leq \epsilon_g$ som termineringsvillkor, där ϵ_F är en skalär och ϵ_g en vektor med toleranser.

- Implementera objektfunktion och/eller bivillkorsfunktioner samt deras jacobianer för lösning av problem (3) med bivillkoren (6) och (7) samt ett av (4) och (5). Välj själv om ni vill lägga modellen i objektfunktionen, dvs

$$F_i(x) = \frac{\begin{bmatrix} A_1^T p_i + a_{14} \\ A_2^T p_i + a_{24} \end{bmatrix}}{A_3^T p_i + a_{34}} - \tilde{q}_i$$

  eller i bivillkoren, dvs

$$F_i(x) = q_i - \tilde{q}_i, \; g_i(x) = q_i - \frac{\begin{bmatrix} A_1^T p_i + a_{14} \\ A_2^T p_i + a_{24} \end{bmatrix}}{A_3^T p_i + a_{34}}$$

  eller motsvarande.

- Implementera bestämningen av startapproximationen beskriven i sektion 3.

- Lös problem (3) med simulerade data med normalfördelade slumpfel $\sigma = 10^{-3}, 10^{-2}, 10^{-1}, 1$ på realistiska parametervärden (utdelas separat). Gör experimentell störningsanalys på de geometriska parametrarna. Hur påverkas störningskänsligheten om vi väljer normering (4) resp. (5)?

  Använd startapproximationen från sektion 3 som x_0. Se till att x_0 är korrekt normerad.

  Funktioner för att konvertera mellan geometriska och algebraiska parametrar utdelas separat.

4.2 Extrauppgift 1

Givet projektionsgeometrin för fokus 1 och fokus 2 kan vi rekonstruera 3d-koordinaten för en patientkula p_i genom att

1. Bestämma skärningen $q_i'^{(1)}$ och $q_i'^{(2)}$ för resp röntgenstråle med planet $z = 0$ genom att lösa

$$\begin{aligned}
\begin{bmatrix} u \\ v \end{bmatrix} &= \frac{\begin{bmatrix} a_{11}x + a_{12}y + a_{13}z + a_{14} \\ a_{21}x + a_{22}y + a_{23}z + a_{24} \end{bmatrix}}{a_{31}x + a_{32}y + a_{33}z + a_{34}} \\
\begin{bmatrix} ua_{31}x + ua_{32}y + ua_{33}z + ua_{34} \\ va_{31}x + va_{32}y + va_{33}z + va_{34} \end{bmatrix} &= \begin{bmatrix} a_{11}x + a_{12}y + a_{13}z + a_{14} \\ a_{21}x + a_{22}y + a_{23}z + a_{24} \end{bmatrix} \\
\begin{bmatrix} ua_{31}x + ua_{32}y + ua_{34} \\ va_{31}x + va_{32}y + va_{34} \end{bmatrix} &= \begin{bmatrix} a_{11}x + a_{12}y + a_{14} \\ a_{21}x + a_{22}y + a_{24} \end{bmatrix} \\
\begin{bmatrix} ua_{31} - a_{11} & ua_{32} - a_{12} \\ va_{31} - a_{21} & va_{32} - a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} a_{14} - ua_{34} \\ a_{24} - va_{34} \end{bmatrix}.
\end{aligned}$$

2. Bestämma mittpunkten på den kortaste linjen mellan 3d-strålarna $f_1 - q_i'^{(1)}$ och $f_2 - q_i'^{(2)}$ genom att lösa

$$\min_{p_i, \alpha^{(1)}, \alpha^{(2)}} \left\| \begin{bmatrix} (f_1 - q_i'^{(1)}) & 0 & I_3 \\ 0 & (f_2 - q_i'^{(2)}) & I_3 \end{bmatrix} \begin{bmatrix} \alpha^{(1)} \\ \alpha^{(2)} \\ p_i \end{bmatrix} - \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \right\| \tag{10}$$

   för varje patientmarkör p_i. Problem (10) är ett linjärt minsta-kvadrat-problem och kan därför lösas direkt.

Simulera ett antal patientmarkörer i hela mätområdet och bestäm experimentellt rekonstruktionsfelet för dessa. Var blir rekonstruktionsfelet minst/störst? Vad händer om vi i stället mäter felet relativt tyngdpunkten hos de rekonstuerade patientmarkörerna?

4.3 Extrauppgift 2

Välj en av följande:

1. Implementera reducerade gradientmetoden med linjesökning för lösning av problem (3). Du behöver implementera Newtons metod för lösning av $g(x_k + Z_k p_z + Y_k p_y) = 0$. Kör på samma problem som i grunduppgiften och jämför konvergenshastighet samt exekveringstid.

2. Antag att vi placerat ett referensobjekt med markörer r_i bredvid patienten och vill använda dessa till att förbättra vår bestämning av projektionsgeometrin för bägge fokus. Hur skulle den nya parametervektorn x se ut? Hur skulle residualfunktionen $F(x)$ och bivillkorsfunktionerna $g_i(x)$ samt deras jacobianer $J(x)$ resp. $A(x)$ se ut? Hur skulle de se ut om vi dessutom kände till att avståndet mellan punkterna $\|r_{i+1} - r_i\| = k$ var konstant?

3. (Förutsätter extrauppgift 1.)

   - Undersök vad som händer om våra data kommer från en projektion som uppfyller bivillkoren (6) och (7), *men vi ej antar det* när vi löser problem (3), dvs generera data från A som uppfyller (6) och (7) och lös (3) utan bivillkoren (6) och (7).
   - Undersök motsatsen, dvs vi antar (6) och (7) uppfyllt, men data kommer från A som ej uppfyller dem.

5 Betyg

Korrekt löst grunduppgift ger betyget 3 på laborationen. Varje extrauppgift 1 grad till. Även kvalitén på lösningen och rapporten kan påverka betyget.

Förutom det som beskrivs ovan ska källkoden ingå i rapporten. Källkoden skall vara väl skriven, indenterad och kommenterad.

Uppgiften skall utföras enskilt och skall inlämnas senast kl. 9.30 den 12 januari 2000.

Samtliga uppgifter skall vara kompletterade senast kl. 09.30 den 23 januari 2000.