Todo list

Part 1: Computed tomography

In this part we will look at a least-squares problem appearing in (a simplified version of) computed tomography (CT). The objective in CT is to determine the structure of a d-dimensional object (typically, a patient) from a series of $d-1$-dimensional X-ray pictures, which are taken from various angles, see Fig. 1.

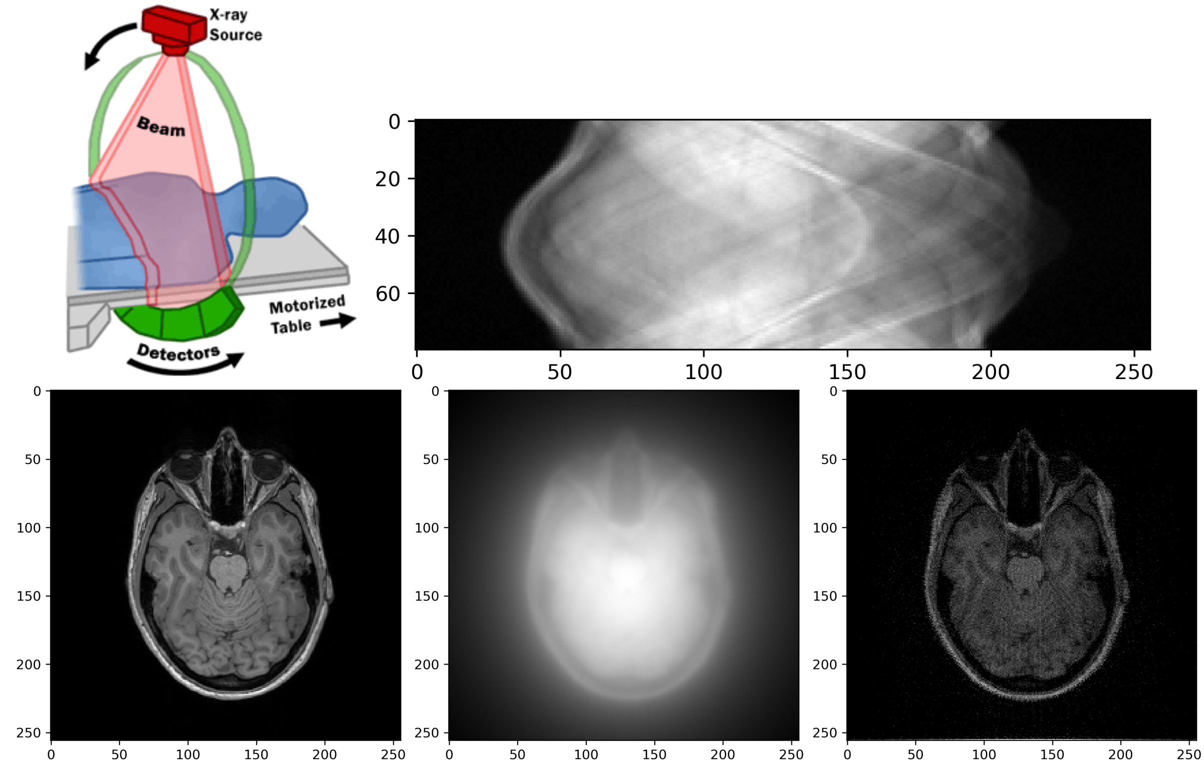


Figure 1: Illustration of the CT process. *Upper row, left-to-right*: drawing of a CT device; CT measurements, where each horizontal line represents one of 80 1-dimensional X-ray pictures of a 2-dimensional object. In this case, measurements contain 1% noise. All measurements are ultimately concatenated in a long vector b. *Lower row, left-to-right*: the unknown 2-dimensional object, which we are trying to reconstruct based on CT measurements; the right-hand side, $A^T b$, in the normal equations; the least squares reconstruction $\hat{x}$, which solves the normal equations $A^T A\hat{x} = A^T b$.

Let us consider a particularly simplified model. We focus on $d = 2$, so that our X-ray pictures are 1-dimensional. We assume that our object is given by $N \times N$ gray-scale pixels, and we take only 4 pictures of it from the angles 0°, 90°, 45°, and 135°, with the resolution of one X-ray per pixel. This situation is outlined in Fig. 2.

1. Based on the description in Fig. 2, find the matrix A representing our measurement model. That is, find A, such that the CT problem to determine $x = [x_1, x_2, x_3, x_4]$ from the measurements (without errors) reduces to solving a linear system $Ax = b$. What is the size of A?

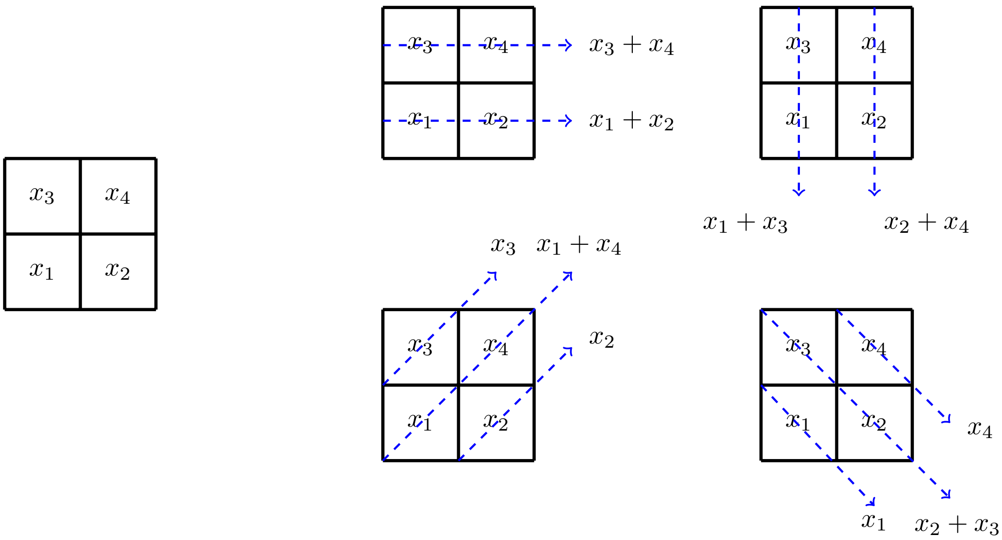


Figure 2: Our CT model with four "X-ray pictures" of a 2×2 object $[x_1, x_2, x_3, x_4] \in \mathbb{R}^4$ from four angles. In this model we measure (assuming that the measurements do not contain errors) a vector $b = [x_1 + x_2, x_3 + x_4, x_1 + x_3, x_2 + x_4, x_2, x_1 + x_4, x_3, x_1, x_2 + x_3, x_4] \in \mathbb{R}^{10}$. In reality, typically we will measure a vector $\tilde{b} = b + f \in \mathbb{R}^{10}$, where $f \in \mathbb{R}^{10}$ is a vector with unknown/random measurement errors.

Answer

$$Ax = b \iff A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_3 + x_4 \\ x_1 + x_3 \\ x_2 + x_4 \\ x_2 \\ x_1 + x_4 \\ x_3 \\ x_1 \\ x_2 + x_3 \\ x_4 \end{bmatrix} \iff \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_3 + x_4 \\ x_1 + x_3 \\ x_2 + x_4 \\ x_2 \\ x_1 + x_4 \\ x_3 \\ x_1 \\ x_2 + x_3 \\ x_4 \end{bmatrix}$$

$A = n \times m$, $x = m \times p$ then $b = n \times p$
The size of A must be $\mathbf{10} \times \mathbf{4}$ matrix because if $A = 10 \times 4$ and $x = 4 \times 1$ then $b = 10 \times 1$.

2. Assuming now that the measurements contain unknown errors, i.e., that instead of the exact vector b we measure a different vector $\tilde{b} = b + f$. Explain why the system $Ax = \tilde{b}$ may lack solutions. Explain the concept "least squares problem" for such systems of linear algebraic equations.

Answer When there is noise in the solution, it may become inconsistent (not solvable), as the measured value is no longer just dependent on the pixels it passes through. It therefore becomes the matter of solving the least squares problem, which finds the best approximate solution, by solving for the x where Ax gives the minimum difference to b. Hence finding the sum where $b - Ax$ is minimized.

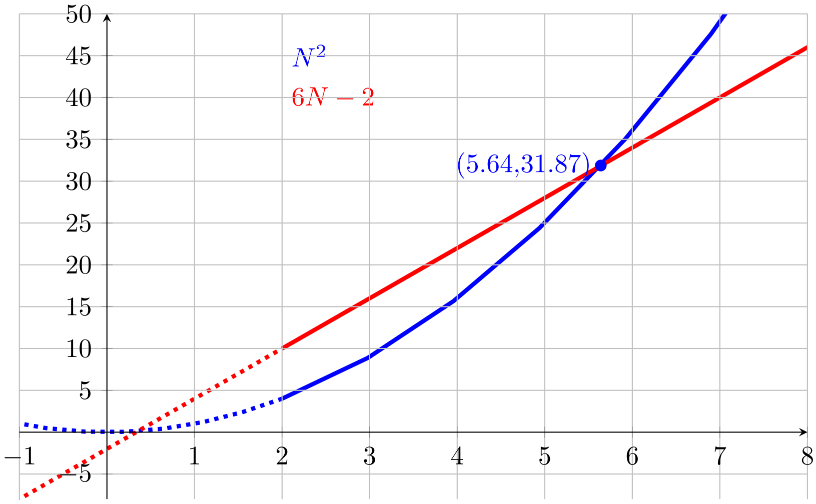
3. Let us now look at an object consisting of $N \times N$ grayscale pixels, where $N \geq 2$. Assuming that we still only take 4 X-ray pictures as in Fig. 2, what are the dimensions of x, b, and A in this situation?

Answer

$$\begin{aligned} A &= 6N - 2 \times N^2 \\ x &= N^2 \times 1 \\ b &= 6N - 2 \times 1 \end{aligned}$$

4. Continuing as in the previous question with a $N \times N$ object, utilize Theorem 8 from Section 1.7 in [Lay] to determine, for which $N \geq 2$ the columns in the matrix A are necessarily *linearly dependent*?

Answer **Definition:** If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v}_1, ..., \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.
We can insert $6n - 1$ and n^2 into a graph and check for which values $N^2 > 6N - 2$.



We can see on the graph that if $N \geq 5.64$, then the columns in the matrix A are necessarily linearly dependent because $N^2 > 6N - 2$ whenever $N > 5.64$. This means that A has more columns than entries whenever $N > 5.64$ and is therefore linearly for $N > 5.64$.

$$n^2 - 6n + 2 = 0$$
$$n_1 = \frac{6 + \sqrt{6^2 - 4 * 1 * 2}}{2 * 1} = 5.64$$
$$n_2 = \frac{6 - \sqrt{6^2 - 4 * 1 * 2}}{2 * 1} = 0.35$$

We choose n_1 because $n \geq 2$.

5. Explain how QR-factorization of a matrix A can be used to solve a least-squares problem $\min_x \|Ax - b\|^2$.

Answer QR-factorization of a matrix allows expressing a matrix A as the product of two matrices Q and R.

Mere fyldestgrende

6. On moodle there are Python and Matlab scripts, illustrating how a QR factorization of a given matrix can be computed.[1] Compute a QR-factorization of the matrix A associated with our simplified CT measurement model for a $N \times N$ object with $N = 2$. Furthermore, using this information compute the orthogonal projection of the vector $b = [0, 1, 2, 3, 4, 5, 6, 7, 8, 9] \in \mathbb{R}^{10}$ onto the linear subspace spanned by the columns of A (recall that $\text{col}(A) = \text{col}(Q)$).

Answer By using MATLAB the orthogonal basis of the vector b is calculated as follows:

[1] They also illustrate, how to multiply matrices. To solve a system of linear algebraic equations, use "numpy.linalg.solve" in Python or "backslash" operator in Matlab.

We start by finding Q and R by using MATLAB.

$$Q = \begin{bmatrix} -0.5000 & -0.3873 & 0.2108 & 0.1782 \\ 0 & 0 & -0.5270 & -0.4454 \\ -0.5000 & 0.1291 & -0.4216 & 0.1782 \\ 0 & -0.5164 & 0.1054 & -0.4454 \\ 0 & -0.5164 & 0.1054 & 0.0891 \\ -0.5000 & 0.1291 & 0.1054 & -0.4454 \\ 0 & 0 & -0.5270 & 0.0891 \\ -0.5000 & 0.1291 & 0.1054 & 0.0891 \\ 0 & -0.5164 & -0.4216 & 0.1782 \\ 0 & 0 & 0 & -0.5345 \end{bmatrix}$$

$$R = \begin{bmatrix} -2 & -0.5 & -0.5 & -0.5 \\ 0 & -1.9365 & -0.3873 & -0.3873 \\ 0 & 0 & -1.8974 & -0.3162 \\ 0 & 0 & 0 & -1.8708 \end{bmatrix}$$

We now have an orthonormal basis Q and the upper triangular matrix R by definition of QR-factorization. By using the formula $Q(Q^T\mathbf{b})$ we can compute the projection:

$$proj_Q b = \begin{bmatrix} 3.5714 \\ 5.5714 \\ 4.2381 \\ 4.9048 \\ 1.9524 \\ 4.5714 \\ 2.6190 \\ 1.6190 \\ 4.5714 \\ 2.9524 \end{bmatrix}$$

7. Find the solution to the least-squares problem associated with our simplified CT model for a $N \times N$ object with $N = 2$ corresponding to the measurements $b = [0, 1, 2, 3, 4, 5, 6, 7, 8, 9] \in \mathbb{R}^{10}$. Use two alternative methods: solving the normal equations and utilizing the QR-factorization.

Answer Method 1 (normal equations):

$$\hat{\mathbf{x}} \iff A^T A\mathbf{x} = A^T\mathbf{b} \iff \begin{bmatrix} 1.6190 \\ 1.9524 \\ 2.6190 \\ 2.9524 \end{bmatrix}$$

Method 2 (Utilizing QR-factorization):

$$\hat{\mathbf{x}} \iff R\mathbf{x} = Q^T\mathbf{b} \begin{bmatrix} 1.6190 \\ 1.9524 \\ 2.6190 \\ 2.9524 \end{bmatrix}$$

Part 2: QR-factorization using Given's rotations

Consider a nonzero vector $[a, b] \in \mathbb{R}^2$. Given's rotation is an orthogonal transformation (as the name suggests, a rotation) which maps $[a, b]$ to a vector $[d, 0] \in \mathbb{R}^2$ aligned with the first coordinate axis.

1. Use Theorem 7 in Subsection 6.2 [Lay], to determine $|d|$ given $[a, b]$.

Answer

$$|d| = Ud = \begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} d \end{bmatrix} = \begin{bmatrix} ad \\ bd \end{bmatrix} = \begin{bmatrix} 1 \\ bd \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \sqrt{1^2 + 0^2} = \sqrt{1}$$

$$||Ux|| = \begin{bmatrix} a \\ b \end{bmatrix} = \sqrt{a^2+b^2} = \sqrt{d^2+0^2} = ||d||$$

Det er forkert

By using the theorem, we have determined $|d|$ as $\sqrt{1}$.

2. Verify that

$$G = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \tag{1}$$

where $c = a/\sqrt{a^2+b^2}$ and $s = b/\sqrt{a^2+b^2}$ is a Given's rotation. That is, check that G is an orthogonal matrix that maps $[a, b]$ to $[d, 0]$.

Answer To check if the G is a orthogonal matrix we can check that $G^TG = I$ is true.

$$G^TG = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} cc+ss & cs-sc \\ sc-cs & ss+cc \end{bmatrix} = \begin{bmatrix} cc+ss & 0 \\ 0 & ss+cc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$G^TG = I$ is in fact true, therefore G is a orthogonal matrix. Now we can check if G maps $[a, b]$ to $[d, 0]$

$$\begin{aligned}\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} &= \begin{bmatrix} a/\sqrt{a^2+b^2} & b/\sqrt{a^2+b^2} \\ -(b/\sqrt{a^2+b^2}) & a/\sqrt{a^2+b^2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a/\sqrt{a^2+b^2} * a + b/\sqrt{a^2+b^2} * b \\ -(b/\sqrt{a^2+b^2}) * a + a/\sqrt{a^2+b^2} * b \end{bmatrix} \\ &= \begin{bmatrix} a^2/\sqrt{a^2+b^2} + b^2/\sqrt{a^2+b^2} \\ 0 \end{bmatrix}\end{aligned}$$

We can see that in fact, G does map $[a, b]$ to a vector $[d, 0]$ in $\mathbb{R}^2$.

3. Let us now consider a vector $x \in \mathbb{R}^m$, $m \geq 2$, such that $x_i = a$ and $x_j = b$, $i < j$. We compute c and s as in the previous question. Let now $G(i, j, a, b)$ be an $m \times m$ matrix, with all rows/columns as in the identity matrix, with the exception of the rows/columns i and j, where we "insert" the matrix G from (1), i.e.: $G(i, j, a, b)_{ii} = G(i, j, a, b)_{jj} = c$, $G(i, j, a, b)_{ij} = -G(i, j, a, b)_{ji} = s$:

$$G(i,j,a,b) = \begin{bmatrix} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{bmatrix}. \tag{2}$$

Verify that $G(i, j, a, b)$ is an orthogonal matrix, and that

$$G(i,j,a,b)x = \begin{bmatrix} x_1 \\ \vdots \\ x_{i-1} \\ d \\ x_{i+1} \\ \vdots \\ x_{j-1} \\ 0 \\ x_{j+1} \\ \vdots \\ x_m \end{bmatrix} \tag{3}$$

Answer

$$GG^T = \begin{bmatrix} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{bmatrix} \begin{bmatrix} I & & & & \\ & c & & -s & \\ & & I & & \\ & s & & c & \\ & & & & I \end{bmatrix} = \begin{bmatrix} I & & & & \\ & cc+ss & & sc-sc & \\ & & I & & \\ & sc-cs & & ss+cc & \\ & & & & I \end{bmatrix} = \begin{bmatrix} I & & & & \\ & 1 & & & \\ & & I & & \\ & & & 1 & \\ & & & & I \end{bmatrix}$$

$$G(2,4,a,b)x = \begin{bmatrix} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{bmatrix} \begin{bmatrix} x_1 \\ a \\ x_3 \\ b \\ x_5 \end{bmatrix} = \begin{bmatrix} x_1 \\ ca+sb \\ x_3 \\ -sa+cb \\ x_5 \end{bmatrix} = \begin{bmatrix} x_1 \\ ca+sb \\ x_3 \\ -(b/\sqrt{a^2+b^2})a + (a/\sqrt{a^2+b^2})b \\ x_5 \end{bmatrix} = \begin{bmatrix} x_1 \\ d \\ x_3 \\ 0 \\ x_5 \end{bmatrix}$$

$$G(i,j,a,b)x = \begin{bmatrix} I & & & & & & \\ & \ddots & & & & & \\ & & c_{ii} & & s_{ij} & & \\ & & & I & & & \\ & & -s_{ji} & & c_{jj} & & \\ & & & & & \ddots & \\ & & & & & & I \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{i-1} \\ a \\ x_{i+1} \\ \vdots \\ x_{j-1} \\ b \\ x_{j+1} \\ \vdots \\ x_m \end{bmatrix}$$

For all $x_1 \dots x_{i-1}, x_{i+1} \dots x_{j-1}$ and $x_{j+1} \dots x_m$ the corresponding rows in G are the identity matrix which means the values will match the original x value.

For x_i the value will be the row i of G multiplied by x. Since the only non-zero values in the row i of G is c and s the value will equal:

$$cx_i + sx_j = ca + sb = a\frac{a}{\sqrt{a^2+b^2}} + b\frac{b}{\sqrt{a^2+b^2}} = \frac{a^2}{\sqrt{a^2+b^2}} + \frac{b^2}{\sqrt{a^2+b^2}} = \frac{a^2+b^2}{\sqrt{a^2+b^2}} = d$$

For x_j the value will be the row j of G multiplied by x. Since the only non-zero values in the row j of G is $-s$ and c, the values will equal:

$$-sx_i + cx_j = -sa + cb = -a\frac{b}{\sqrt{a^2+b^2}} + b\frac{a}{\sqrt{a^2+b^2}} = \frac{-ab}{\sqrt{a^2+b^2}} + \frac{ab}{\sqrt{a^2+b^2}} = \frac{-ab+ab}{\sqrt{a^2+b^2}} = 0$$

4. Explain why a product of square orthogonal matrices is itself an orthogonal matrix. I.e., assuming that $Q_1, Q_2, \dots, Q_k$ are orthogonal matrices, explain why the matrix $Q_1 Q_2 \cdots Q_k$ is orthogonal.

Answer Let A and B be orthogonal matrices.

$$AA^T = A^T A = I$$

and

$$BB^T = B^T B = I$$

Then we have a new vector AB. An matrix is a orthogonal matrix if $G^T G = I$ is true. Let's try to do that by replacing G with our new vector AB and check if it equals the identity matrix.

$$(AB)^T(AB) = (B^T A^T)AB = \mathbf{B^T(A^T A)B} = B^T(I)B = B^T B = I$$

The equation highlighted in **BOLD** are using the matrix property $(AB)C = A(BC)$.

We can see that the product of two orthogonal matrices is a new orthorgnal matrix because the rule $G^T G = I$ is true. Hence we get a new orthogonal matrix. Since they are all orthogonal matrices, the matrix $Q_1 Q_2 \cdots Q_k$ is orthogonal.

5. Consider the following algorithm, where the matrix A of size $m \times n$ is given. *Dette er noget sygt fyldtekst*

$Q := I_{m \times m}$; $R := A$;
for $i = 1, \dots, n$ **do**
  for $j = i+1, \dots, m$ **do**
    $a = R_{ii}$; $b = R_{ji}$;
    if $b \neq 0$ **then**
      $R := G(i,j,a,b)R$; $Q := QG(i,j,a,b)^T$;
    end if
  end for
end for

where $G(i,j,a,b)$ is given by (2).

(a) Use the previous questions to explain, why the matrix Q remains orthogonal throughout all iterations of the algorithm.

Answer Mangler 5a

(b) Explain why the equation $A = QR$ is fullfilled throughout all iterations of the algorithm.

Answer Mangler 5b

(c) Use (3) to explain, why after the algorithm terminates, the computed matrix R is upper triangular.

Answer Mangler 5c