Math 415 - Lecture 27

An application of QR-decomposition, Change of basis

Friday October 30th 2015

Textbook reading: Chapter 3.4, Chapter 2.6

Suggested practice exercises: Chapter 2.6: Exercises 36, 37, 38,39, 40,43

Khan Academy video: Change of basis

Strang lecture: Change of basis; image compression

1 Review

Theorem 1 (QR decomposition)**.** *Let A be a $m \times n$ matrix of rank n with linear independent columns. There is an orthogonal matrix $m \times n$-matrix Q and an upper triangular $n \times n$ invertible matrix R such that*

$$A = QR.$$

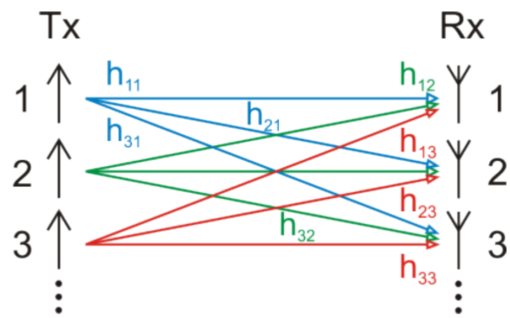
Theorem 2. *Let A be a matrix with linear independent columns. Suppose $A\mathbf{x} = \mathbf{b}$ has no solution. Then the solution of $R\mathbf{x} = Q^T\mathbf{b}$ is the least square solution of $A\mathbf{x} = \mathbf{b}$.*

2 An application in wireless communication

In multiple-input and multiple-output (short: MIMO) systems, a transmitter sends multiple streams by multiple transmit antennas. Let us suppose there are n transmitters and m receivers. This can modelled using Linear Algebra:

$$\underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{\text{received vector } \mathbf{y}} = \underbrace{\begin{bmatrix} h_{1,1} & \dots & h_{1,n} \\ \vdots & \ddots & \vdots \\ h_{m,1} & \dots & h_{m,n} \end{bmatrix}}_{\text{channel matrix } H} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{\text{transmitted vector } \mathbf{x}} .$$

Suppose that the channel matrix H is known both to person A who sending information and to person B who is receiving the information.



Let us try and understand the engineering meaning of some of the linear algebra of the matrix H and the equation $y = Hx$. Remember: the x vector describes what the transmiter is sending out and y is the vector describing what is received.

We want to understand

$$y = Hx.$$

- What is the first column of H?
- What is Nul(H)? If the signal x belongs to the nullspace, what signal y will be received?
- In a well designed system you want Dim(Nul(H)) =?
- What is $Col(H)$?

When B receives the signal, she wants to reconstruct the vector $\mathbf{x}$. Optimally, she would just solve the linear system

$$H\mathbf{x} = \mathbf{y}.$$

Unfortunately, almost always B received $\mathbf{y} + \epsilon$ instead of $\mathbf{y}$, where $\epsilon \in \mathbb{R}^m$ is noise.

So B would try to solve

$$H\mathbf{x} = \mathbf{y} + \epsilon.$$

instead. However, that system might not have a solution. So B has to find the least square solution! Because B receives many messages from A, she will have to find the least square solution many times. Luckily, H does not change, and has independent columns (Nul(H) = 0). So B determines the QR-decomposition of H

$$H = QR,$$

once, and then just solves

$$R\mathbf{x} = Q^T(\mathbf{y} + \epsilon)$$

each time she receives a new message. This is easy to do, since R is upper triangular.

3 Linear transformation revisited

Recall the notion of coordinate vectors. If $\mathcal{B} = \{b_1, b_2\}$ is a basis for $\mathbb{R}^2$, and x some vector then the coordinate vector of x is $x_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ precisely if $x = c_1 b_1 + c_2 b_2$. We want to understand how to relate coordinate vectors $x_{\mathcal{B}}$ and $x_{\mathcal{C}}$ for different bases $\mathcal{B}$ and $\mathcal{C}$. We will see that there is for every two bases a matrix $I_{\mathcal{C},\mathcal{B}}$ so that

$$x_{\mathcal{C}} = I_{\mathcal{C},\mathcal{B}} x_{\mathcal{B}}.$$

Remember Theorem 1 of Lecture 17? Here it is again.

Theorem 3. *Let $\mathcal{B}$ be a basis of $\mathbb{R}^m$ and $\mathcal{C}$ be a basis of $\mathbb{R}^n$ and let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. Then there is a $n \times m$ matrix $T_{\mathcal{C},\mathcal{B}}$ such that for every $\mathbf{v} \in \mathbb{R}^m$*

$$T(\mathbf{v})_{\mathcal{C}} = T_{\mathcal{C},\mathcal{B}} \mathbf{v}_{\mathcal{B}}.$$

and

$$T_{\mathcal{C},\mathcal{B}} = \begin{bmatrix} T(\mathbf{v_1})_{\mathcal{C}} & T(\mathbf{v_2})_{\mathcal{C}} & \dots & T(\mathbf{v_m})_{\mathcal{C}} \end{bmatrix}$$

where $\mathcal{B} = \left(\mathbf{v_1}; \dots; \mathbf{v_m}\right)$.

We will use this first in the special case $T = I$, where $I(v) = v$ (seemingly boring!).

Example 4. Consider $\mathcal{E} := \{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$ and $\mathcal{B} := \{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$. Let $I : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation (the Identity!)

$$I(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Find the matrix $I_{\mathcal{E},\mathcal{B}}$ that represents I with respect to the input basis $\mathcal{B}$ and output basis $\mathcal{E}$.

Solution. By definition the matrix $I_{\mathcal{E},\mathcal{B}}$ has as first column b_1 expressed in the standard basis, and as second column b_2 also expressed in the standard basis. But for any vector $x \in \mathbb{R}^n$ we have $x_{\mathcal{E}} = x$! So

$$I_{\mathcal{E},\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} b_1 & b_2 \end{bmatrix}.$$

Example 5. Given $\mathbf{v} \in \mathbb{R}^2$ what is $I_{\mathcal{E},\mathcal{B}} \mathbf{v}_{\mathcal{B}}$?

Solution. Let $\mathbf{v}_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$. Then

$$I_{\mathcal{E},\mathcal{B}} \mathbf{v}_{\mathcal{B}} = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = c_1 b_1 + c_2 b_2 = v!$$

Suppose $\mathbf{v}$ is a vector in $\mathbb{R}^n$, and we have two bases in $\mathbb{R}^n$. so that we get two coordinate vectors $\mathbf{v}_{\mathcal{C}}$ and $\mathbf{v}_{\mathcal{B}}$. How are they related?

Theorem 6. *Let $\mathcal{B}$ be a basis of $\mathbb{R}^n$ and $\mathcal{C}$ be another basis of $\mathbb{R}^n$ and let $I : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation such that $I(\mathbf{v}) = \mathbf{v}$ for every $\mathbf{v} \in \mathbb{R}^n$. Then*

$$\mathbf{v}_{\mathcal{C}} = I_{\mathcal{C},\mathcal{B}}\mathbf{v}_{\mathcal{B}}.$$

We call the matrix $I_{\mathcal{C},\mathcal{B}}$ a **change of base matrix**, it transforms coordinate vectors from the $\mathcal{B}$ to the $\mathcal{C}$ basis.

Example 7. Let $\mathcal{E}$ be the standard basis of $\mathbb{R}^n$ and $\mathcal{B}$ be another basis of $\mathbb{R}^n$. What is $I_{\mathcal{E},\mathcal{B}}$?

Solution. The columns of $I_{\mathcal{E},\mathcal{B}}$ are the basic vectors $b_1, b_2, \dots$ expressed in the standard basis. So

$$I_{\mathcal{E},\mathcal{B}} = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}$$

So this is the *easy* change of basis matrix: you just write down the $\mathcal{B}$ basis as columns of your matrix. It has the property that

$$v = v_{\mathcal{E}} = I_{\mathcal{E},\mathcal{B}} v_{\mathcal{B}}$$

Example 8. Let $\mathcal{B}$ be a basis of $\mathbb{R}^n$ and $\mathcal{C}$ be a basis of $\mathbb{R}^n$. What is $I^{-1}_{\mathcal{C},\mathcal{B}}$?

Solution. $I_{\mathcal{C},\mathcal{B}}$ is the matrix with columns the $\mathcal{B}$ basis vectors expressed in the $\mathcal{C}$ basis, and $I^{-1}_{\mathcal{C},\mathcal{B}}$ is the inverse of this matrix. These matrices have the property that

$$v_{\mathcal{C}} = I_{\mathcal{C},\mathcal{B}} v_{\mathcal{B}}, \quad v_{\mathcal{B}} = I^{-1}_{\mathcal{C},\mathcal{B}} v_{\mathcal{C}}.$$

Example 9. As before, let $\mathcal{E} := \{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$ and $\mathcal{B} := \{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$. What is $I_{\mathcal{B},\mathcal{E}}$?

Solution. We know what $I_{\mathcal{E},\mathcal{B}}$ is, it is just $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Then $I_{\mathcal{B},\mathcal{E}}$ is the transition matrix going the other way, so it is the inverse of the *easy* matrix, so

$$I_{\mathcal{B},\mathcal{E}} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Example 10. Let $\mathcal{E}$ be the standard basis of $\mathbb{R}^n$ and $\mathcal{C}$ be a orthonormal basis of $\mathbb{R}^n$. Then $I_{\mathcal{C},\mathcal{E}} = I^T_{\mathcal{E},\mathcal{C}}$. Why?

Solution. $I_{\mathcal{E},\mathcal{C}}$ the matrix with orthonormal columns, so it is an orthogonal matrix. $I_{\mathcal{C},\mathcal{E}}$ is the inverse. But the inverse of an orthogonal matrix is easy, just the transpose.

Theorem 1. *Let $\mathcal{B} := (\mathbf{u_1}, \dots, \mathbf{u_n})$ be an orthonormal basis of $\mathbb{R}^n$ and $U = \begin{bmatrix} \mathbf{u_1} & \dots & \mathbf{u_n} \end{bmatrix}$. Then for every $\mathbf{v} \in \mathbb{R}^n$*

$$v_{\mathcal{B}} = U^T v.$$

4 Change of basis

Theorem 11. *Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation and $\mathcal{A}$ and $\mathcal{B}$ be two bases of $\mathbb{R}^m$ and $\mathcal{C}, \mathcal{D}$ be two bases of $\mathbb{R}^n$. Then*

$$T_{\mathcal{C},\mathcal{A}} = I_{\mathcal{C},\mathcal{D}} T_{\mathcal{D},\mathcal{B}} I_{\mathcal{B},\mathcal{A}}.$$

$$\begin{tikzcd}
(\mathbb{R}^m, \mathcal{A}) \arrow[r, "\text{apply } T_{\mathcal{C},\mathcal{A}}"] \arrow[d, "I_{\mathcal{B},\mathcal{A}}"'] & (\mathbb{R}^n, \mathcal{C}) \\
(\mathbb{R}^m, \mathcal{B}) \arrow[r, "\text{apply } T_{\mathcal{D},\mathcal{B}}"] & (\mathbb{R}^n, \mathcal{D}) \arrow[u, "I_{\mathcal{C},\mathcal{D}}"']
\end{tikzcd}$$

Example 12. Consider $\mathcal{B} := \mathcal{D} := \{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$ and $\mathcal{A} := \mathcal{C} := \{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ as before. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be again the linear transformation that

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto A \begin{bmatrix} x \\ y \end{bmatrix}, \text{ where } A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Determine $T_{\mathcal{C},\mathcal{C}}$.

Solution. By Theorem 11

$$T_{\mathcal{C},\mathcal{C}} = I_{\mathcal{C},\mathcal{B}} T_{\mathcal{B},\mathcal{B}} I_{\mathcal{B},\mathcal{C}}.$$

In Lecture 27, we already calculated that

$$I_{\mathcal{C},\mathcal{B}} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, I_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Since $\mathcal{B}$ is the standard basis,

$$T_{\mathcal{B},\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Therefore

$$T_{\mathcal{C},\mathcal{C}} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 1 & 5 \end{bmatrix}$$

.

Example 13. Let $\mathcal{E}$ be the standard basis of $\mathbb{R}^n$, let $\mathcal{B} := (\mathbf{u_1}, \dots, \mathbf{u_n})$ be an orthonormal basis of $\mathbb{R}^n$ and $U = \begin{bmatrix} \mathbf{u_1} & \dots & \mathbf{u_n} \end{bmatrix}$ Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Then

$$T_{\mathcal{B},\mathcal{B}} = U^T T_{\mathcal{E},\mathcal{E}} U.$$

Why?

Solution.