Problem 1: Fixed point iterations in practice

Suppose we want to find a solution for:

$$x^4 - x - 10 = 0.$$

Notice that the solutions for this polynomial are $x_1 \approx 1.85556$ and $x_2 = -1.6975$.

1. Plot this function (either by hand or with a software of your choice).

2. Consider the following fixed point scheme:

$$x_{k+1} = \frac{10}{x_k^3 - 1}.$$

   Let $x_0 = 2$. Does this method converge? Why (/why not)? In the same plot draw the functions $y = x$, $y = \frac{10}{x^3-1}$ and add to this plot the first 5 iterations of such method.

3. Consider the following fixed point scheme:

$$x_{k+1} = (x_k + 10)^{1/4}.$$

   Test this method with the following starting points: $x_0 = 1$, $x_0 = 2$, $x_0 = 4$. Does this method converge? Why (/why not)? In the same plot draw the functions $y = x$, $y = (x + 10)^{1/4}$ and add to this plot the first 5 iterations of such method for $x_0 = 2$.

4. Consider the following fixed point scheme:

$$x_{k+1} = \frac{(x_k + 10)^{1/2}}{x_k}.$$

   Test this method with $x_0 = 1.8$, $x_0 = 2$. Does this method converge? Why (/why not)? In the same plot draw the functions $y = x$, $y = \frac{(x+10)^{1/2}}{x}$ and add to this plot the first 5 iterations of such method for $x_0 = 1.8$.

What can you tell about the methods that converge? How do these methods look? Can you relate this to the contraction mapping theorem? What can you tell about the speed of convergence?

Convergence of iterative methods

Consider the following theorem:

Theorem 1. *Let $f \in C^2$ on $I_\delta = [\xi - \delta, \xi + \delta], \delta > 0$. Assume that $f(\xi) = 0$ and $f''(\xi) \neq 0$. Assume that $\exists A > 0$ such that $\frac{|f''(x)|}{|f'(y)|} \leq A$ for all $x, y \in I_\delta$. If x_0 is such that $|\xi - x_0| \leq min(\delta, \frac{1}{A})$, then $x_n \to \xi$ quadratically.*

1. Consider the function $f(x) = sin(x)$ ~~show~~ that if $x_0 \in (\frac{-\pi}{2} + a, \frac{\pi}{2} - a)$, where $a \geq 0$ and $x_0 \neq 0$, $x_n \to 0$ quadratically.
2. What assumptions do you need to have on a?

Newton's Method

1. Suppose Newton's method is applied to find the solution of the following equation:

$$e^x - x - 2 = 0.$$

Write down Newton's iteration. Show that if the starting value is positive, the iteration converges to the positive solution, and if the starting value is negative, it converges to the negative solution. Obtain approximate expressions for x_1 if we start in $x_0 = 100$ and if we start in $x_0 == -100$, describe the subsequent behaviour of the iteration.

2. Consider the iteration for the solution $f(x) = 0$:

$$x_{k+1} = x_k - \frac{(f(x_k))^2}{f(x_k + f(x_k)) - f(x_k)}.$$

Explain the connection with Newton's method, and show that it converges quadratically if x_0 is sufficiently close to the solution.

Secant method

1. Show that the secant method,

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k), \tag{1}$$

can be written as:

$$x_{k+1} = \frac{x_k f(x_{k-1}) - x_{k-1} f(x_k)}{f(x_{k-1}) - f(x_k)}.$$

■ Show that minimizing a function $f: \mathbb{R} \to \mathbb{R}$ using Newton's method is equivalent to minimizing a sequence of quadratics.

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)} \Rightarrow \boxed{x_{n+1} - x_n = -\frac{f'(x_n)}{f''(x_n)}}$$

$$f(x_{n+1}) = f(\underbrace{x_{n+1} - x_n}_{\Delta x_n} + x_n)$$

$$f(x_n + \Delta x_n) = f(x_n) + f'(x_n)\Delta x_n + \frac{f''(x_n)(\Delta x_n)^2}{2} + O((\Delta x_n)^3)$$

$$\Delta x_n = x_{n+1} - x_n \Rightarrow x_n + \Delta x_n = x_{n+1}$$

$$f(x_{n+1}) = f(x_n) + f'(x_n)(x_{n+1} - x_n) + \frac{f''(x_n)(x_{n+1} - x_n)^2}{2} + O((\Delta x_n)^2)$$

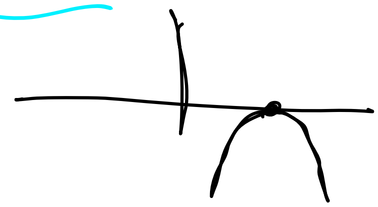
$$f(x_{n+1}) = f(x_n) + f'(x_n)\left(-\frac{f'(x_n)}{f''(x_n)}\right) + \frac{1}{2} f''(x_n)\left(-\frac{f'(x_n)}{f''(x_n)}\right)^2 + O((\Delta x_n)^2)$$

$$f(x_{n+1}) - f(x_n) = -\frac{(f'(x_n))^2}{f''(x_n)} + \frac{1}{2}\frac{(f'(x_n))^2}{f''(x_n)} + O((\Delta x_n)^2)$$

$$f(x_{n+1}) - f(x_n) = -\frac{1}{2}\frac{\overset{+}{(f'(x_n))^2}}{f''(x_n)} + O((\Delta x_n)^2)$$

$$\leq 0$$

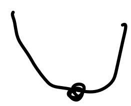
Min $f(x)$



$$\arg\min f(x) = x^* \Rightarrow f'(x^*) = 0$$

$$\& \quad f''(x^*) \geq 0$$

sequence of quadratics $\underbrace{\frac{1}{2}\frac{(f'(x_n))^2}{f''(x_n)}}_{\min = 0}$



$$\Rightarrow f(x_{n+1}) = f(x_n) \Rightarrow x_n = x^*$$

(Minimum unique)

2. Now, denote the root of f to be ξ, so that $f(\xi) = 0$. Also assume that f is twice continuously differentiable and that $f' > 0$ and $f'' > 0$ in a neighborhood of ξ. Define the quantity φ to be:

$$\varphi(x_k, x_{k-1}) = \frac{x_{k+1} - \xi}{(x_k - \xi)(x_{k-1} - \xi)},$$

where x_{k+1} is as in 1. Compute (for fixed value of x_{k-1})

$$\psi(x_{k-1}) = \lim_{x_k \to \xi} \varphi(x_k, x_{k-1}).$$

3. Now compute

$$\lim_{x_{k-1} \to \xi} \psi(x_{k-1}),$$

and therefore show that

$$\lim_{x_k, x_{k-1} \to \xi} \varphi(x_k, x_{k-1}) = \frac{f''(\xi)}{2f'(\xi)}.$$

4. Next, assume that the secant method has convergence order q, that is to say that

$$\lim_{k \to \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^q} = A < \infty.$$

Using the above results, show that $q - 1 - 1/q = 0$, and therefore that $q = (1 + \sqrt{5})/2$.

5. Finally, show that this implies that

$$\lim_{k \to \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^q} = \left(\frac{f''(\xi)}{2f'(\xi)} \right)^{q/(1+q)}.$$

Errors!

1. State the 4 causes of error in numerical schemes and give an example of each one of them.

2. Consider the following approximation to derivatives:

$$f'(x) \approx \frac{3f(x+h) - f(x) - 2f(x-h)}{5h} \quad \checkmark$$

$$f'(x) = \frac{f(x-2h) - 8f(x-h) + 8f(x+h) - f(x+2h)}{12h} + \frac{4h^4}{5!} f^{(5)}(\xi)$$

For the first expression show that its error is of $\boxed{\text{order } \mathcal{O}(h)}$ for the second expression do Taylor expansions and find the error given. We know that $\xi \in [x - 2h, x + 2h]$.

$$f'(x) \approx \frac{3f(x+h) - f(x) - 2f(x-h)}{5h}$$

$$f(x+h) = f(x) + f'(x)h + \frac{f''(\xi_1)h^2}{2}$$

$\xi_1 \in [x, x+h]$

$\neq$

$$\Rightarrow$$

$$3f(x+h) = 3f(x) + 3f'(x)h + \frac{3f''(\xi_1)h^2}{2} \quad A$$

$$f(x-h) = f(x) - f'(x)h + \frac{f''(\xi_2)h^2}{2}$$

$\xi_2 \in [x-h, x]$

$$-2f(x-h) = -2f(x) + 2f'(x)h - f''(\xi_2)h^2 \quad B$$

$$3f(x+h) - 2f(x-h) = f(x) + 5f'(x)h + \underbrace{O(h^2)}_{A \,\&\, B}$$

$[x-h, x+h]$

$$\Rightarrow \boxed{\frac{3f(x+h) - f(x) - 2f(x-h)}{h} + O(h) = f'(x)}$$

3. Consider the forward difference approximation to the derivative:

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}.$$

What happens if $f''(x) = 0$?

4. Consider the central difference approximation to the derivative:

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h},$$

what happens if $f'''(x) = 0$?

5. Consider the following polynomial:

$$p(x) = ax^2 + bx + c.$$

Show that the central difference scheme is "exact" i.e. there is no error.

6. For which type of polynomial is the forward difference approximation to the derivative exact?

7. Consider the following approximation to the integral of f on an interval $[a, b]$:

$$\int_a^b f(x)dx \approx f\left(\frac{a+b}{2}\right)(b-a).$$

Suppose that $b - a = \mathcal{O}(h)$. Derive an error estimate for this approximation using a Taylor expansion. What assumption(s) do you need to take into consideration?

8. Give three examples of functions (different than constants) whose integral on the interval $[0, 2]$ can be exactly calculated using 7. What do these functions have in common?

9. Come up with a Monte-Carlo like method to approximate π. How could you make this method "better"?

Computer representation of numbers

10. If we have "a lot" of computations to do, what would be the problem if we use SP?

Numerical Linear Algebra (so far)

Practice "modelling" with Linear Algebra

1. Suppose that we have a system that can be in 3 different states. If we know that at time t we are in state i, then the probability of going to state j at time $t+1$ is given as:

$$P(X_{t+1} = j | X_t = i) = \begin{cases} 1/2 & \text{if } i+1 = j \text{ and } i < 3 \\ 1/2 & \text{if } i = j \\ 1/2 & \text{if } i = 3 \text{ and } j = 1 \\ 0 & \text{otherwise} \end{cases}.$$

   Build a matrix such that its ij entry represents the probability $P(X_{t+1} = j | X_t = i)$.

2. Find (α_1, α_2) such that $||\vec{x} - \alpha_1 \vec{a_1} - \alpha_2 \vec{a_2}||^2$ is minimized.

Linear Algebra in a computer

1. Write in python-style pseudo-code, an algorithm for multiplying the matrices $a \in R^{m \times p}, b \in R^{p \times n}$. How many floating point adds and multiplies do you need, in terms of m,n and p?

2. Let $B \in \mathbb{R}^{m \times n}$, $A \in \mathbb{R}^{n \times n}$, $u \in \mathbb{R}^{n \times 1}$, $v^\top \in \mathbb{R}^{i \times p}$, and $z \in \mathbb{R}^{p \times 1}$, which of the following options need less floating point adds and multiplies? Why?

   - $(BA)(uv^\top)z$
   - $B(Au)(v \top z)$

3. In numerical analysis inverting a matrix is seen as "terrible". But why? For an invertible $n \times n$ matrix, how many FLOPs would we need to invert it? If you're not quite sure try inverting this matrix by hand and count how many operations you needed.

$$A = \begin{bmatrix} 4 & -2 & 1 \\ 5 & 0 & 3 \\ -1 & 2 & 6 \end{bmatrix}.$$

LU factorization

1. Count FLOPs for forward or backward substition.

2. Prove that the product of two lower unit triangular matrices is unit lower triangular.

$= 0$ if $x \in \mathbb{R}^2$ & a_1 is l.i of a_2

2. Find (α_1, α_2) such that $||\vec{x} - \alpha_1 \vec{a_1} - \alpha_2 \vec{a_2}||_2^2$ is minimized.

$\|x\|_2^2$ if $a_1 = a_2 = 0$

$$\|x - \alpha_1 a_1 - \alpha_2 a_2\|_2^2 =$$

$$\left(x - \alpha_1 a_1 - \alpha_2 a_2 \right)^T \left(x - \alpha_1 a_1 - \alpha_2 a_2 \right)$$

$$= \left(x^T - \alpha_1 a_1^T - \alpha_2 a_2^T \right)\left(x - \alpha_1 a_1 - \alpha_2 a_2 \right)$$

$$= x^T x - \alpha_1 x^T a_1 - \alpha_2 x^T a_2 - \alpha_1 a_1^T x + \boxed{\alpha_1^2 a_1^T a_1} + \alpha_1\alpha_2 a_1^T a_2 - \alpha_2 a_2^T x + \alpha_1 \alpha_2 a_2^T a_1 + \boxed{\alpha_2^2 a_2^T a_2}$$

$$= \alpha_1 \left(- x^T a_1 - x^T a_1 \right) + \alpha_2 \left(- x^T a_2 - x^T a_2 \right) + \alpha_1 \alpha_2 \left(a_1^T a_2 + a_1^T a_2 \right) + \alpha_1^2 a_1^T a_1 + \alpha_2^2 a_2^T a_2 + x^T x$$

$$= \alpha_1 \left(-2 x^T a_1 \right) + \alpha_2 \left(-2 x^T a_2 \right) + \alpha_1 \alpha_2 \left(2 a_1^T a_2 \right) + \alpha_1^2 \left(a_1^T a_1 \right) + \alpha_2^2 \left(a_2^T a_2 \right) + \cancel{\|x\|_2^2}$$

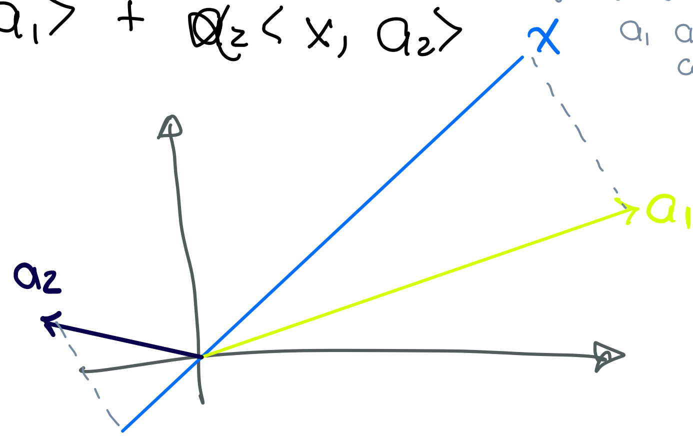
$\|a_1\|_2^2$ $\|a_2\|_2^2$

$$= \alpha_1 \left(-2 \langle x, a_1 \rangle \right) + \alpha_2 \left(-2 \langle x, a_2 \rangle \right) + \alpha_1 \alpha_2 \left(\langle a_1, a_2 \rangle \right) + \alpha_1^2 \|a_1\|_2^2 + \alpha_2^2 \|a_2\|_2^2 + \|x\|_2^2$$

$$= -2 \langle x, \alpha_1 a_1 \rangle - 2 \langle x, \alpha_2 a_2 \rangle + \langle \alpha_1 a_1, \alpha_2 a_2 \rangle + \|\alpha_1 a_1\|_2^2 + \|\alpha_2 a_2\|_2^2 + \|x\|_2^2$$

The α's are the components of the projection of a onto a_1 and a_2 corresp.

$$x = \alpha_1 \langle x, a_1 \rangle + \alpha_2 \langle x, a_2 \rangle$$



if $\mathbb{R}^n = \mathbb{R}^2$ & a_1 l.i a_2

If a_1 and a_2 are linearly independent $\Rightarrow$ they span a plane (of dim 2). Then you're looking for the orthogonal projection of x on that plane. The α's are the components of such projection (think- Gram-Schmidt). Otherwise they span a line and we're looking for the orthogonal projection of $\vec{x}$ on to that line (the α's are not unique).

$\hookrightarrow a_1 = k a_2$

3. For the LU factorization, if we label the k^{th} Gauss transform M_k where $M_K = I - \tau^k e_k^T$, show that $M_k^{-1} = I + \tau^k e_k^T$.

4. Prove that $M_1^{-1}...M_{n-1}^{-1}$ is a lower triangular matrix L.

5. Describe how you can create a recursive algorithm for computing $A = LDU$, where L is unit lower triangular and U is unit upper triangular and D is diagonal (actually diagonal). To make it easy, assume n (number: $A \in \mathbb{R}^{n\times n}$) is a power of two. Hint: what is the recursive base case (n=1)? Think about block matrices.

Norms

1. Let $v \in \mathbb{R}^n$ prove the following inequalities:

$$\|v\|_\infty \leq \|v\|_2$$
$$\|v\|_2^2 \leq \|v\|_1 \|v\|_\infty.$$

2. For each of the previous inequalities, give an example of a non-zero vector in $\mathbb{R}^n$ for which equality is obtained.

3. Show that:

$$\|v\|_\infty \leq \|v\|_2 \leq \|v\|_1$$
$$\|v\|_2 \leq \sqrt{n}\|v\|_\infty$$

4. Let $A \in \mathbb{R}^{m\times n}$ be any real-values matrix. Show that:

$$\|A\|_\infty \leq \sqrt{n}\|A\|_2, \|A\|_2 \leq \sqrt{m}\|A\|_\infty.$$

5. The Frobenius norm for a matrix $A \in \mathbb{R}^{m\times n}$ as:

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

Prove that $\|A\|_F = \sqrt{\text{trace}\,(A^\top A)}$.

6. Prove that $\|A\|_F = \sqrt{\sum_j \sigma_j}$, where σ_j are the singular values of A.

7. Show that $\|A\|_2 \leq \|A\|_F$.

3. For the LU factorization, if we label the k^{th} Gauss transform M_k where $M_K = I - \tau^k e_k^T$, show that $M_k^{-1} = I + \tau^k e_k^T$.

4. Prove that $M_1^{-1}...M_{n-1}^{-1}$ is a lower triangular matrix L.

5. Describe how you can create a recursive algorithm for computing $A = LDU$, where L is unit lower triangular and U is unit upper triangular and D is diagonal (actually diagonal). To make it easy, assume n (number: $A \in \mathbb{R}^{n\times n}$) is a power of two. Hint: what is the recursive base case (n=1)? Think about block matrices.

3 $M_K M_K^{-1} = I = M_K^{-1} M_K$ (if you're asked a matrix is the inverse of anotherone just multiply it!)

$\underset{n\times n}{M_K} M_K^{-1}$

$= (\underset{n\times n}{I} - \underbrace{\underset{n\times 1}{\tau^K} \underset{1\times n}{e_K^T}}_{n\times n})(I + \tau^K e_K^T) = I + I\tau^K e_K^T - I\tau^K e_K^T + (\tau^K e_K^T)(\tau^K e_K^T)$

$= I + \tau^K e_K^T - \tau^K e_K^T + (\tau^K e_K^T)(\tau^K e_K^T)$

$= I + \underset{n\times 1}{\tau^K} \underbrace{(\underset{1\times n}{e_K^T} \underset{n\times 1}{\tau^K})}_{=0} \underset{1\times n}{e_K^T} = I$

$e_K^T \tau^K \in \mathbb{R}$

$= \langle e_K, \tau^K \rangle$

$= [0, 0, \dots, \underset{K\text{-th}}{1}, 0, \dots, 0] \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \tau_{K+1} \\ \vdots \\ \tau_n \end{bmatrix} = 0$

$M_K^{-1} M_K$

$= (I + \tau^K e_K^T)(I - \tau^K e_K^T)$

$= I - \tau^K e_K^T + \tau^K e_K^T - \tau^K \underbrace{(e_K^T \tau^K)}_{=0} e_K^T$

$= I$

4. $M_1^{-1} \cdots M_{n-1}^{-1}$

$M_K^{-1} = I + \tau^K e_K^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \tau_{K+1} \\ \vdots \\ \tau_n \end{bmatrix} [0, 0, \dots, \underset{K\text{-th element}}{1}, 0, \dots, 0]$

lower triangular

$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & \tau_{K+1} & \cdots & 0 \\ \vdots & & & & \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$

K+1, K-th element

0 vector K-th column

$[0 \cdots \tau^K \cdots 0]$

$$\underset{n \times n}{A} = \begin{array}{c c} n & \left[\begin{array}{c|c} a_{11} & v^T \\ \hline u & A_1 \end{array}\right] \\ & n \end{array}$$

where

$a_{11} \in \mathbb{R}$

$v \in \mathbb{R}^{n-1}$

$u \in \mathbb{R}^{n-1}$

$A_1 \in \mathbb{R}^{(n-1)\times(n-1)}$

Cholesky and QR

1. Show that if A is positive definite, then A^k is also positive definite for all $k \geq 1$.

2. Find the QR factorization of the following matrix:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. Let A be an $m \times n$ matrix with independent columns. If $A = QR$ and $A = \hat{Q}\hat{R}$ are QR-factorization of A, show that $Q = \hat{Q}$ and $R = \hat{R}$.

4. Let Q be an orthogonal matrix, prove that for any other $n \times n$ matrix A we have that $\|AQ\|_2 = \|A\|_2$.

5. Let Q be an orthogonal matrix, prove that for any other orthogonal matrix U, QU is also an orthogonal matrix.

6. Prove that for any real matrix $A \in \mathbb{R}^{m\times n}$ with $m \geq n$ $A^\top A$ is symmetric and if $A^\top A$ is nonsingular, then it is positive definite.

7. Consider the following lemma:

   Lemma 1. *Let $A \in \mathbb{R}^{m\times m}$ be spd (symmetric positive definite matrix). Let $X \in \mathbb{R}^{m\times n}$ where $m \geq n$ be full rank (column rank because $m \geq n$). Then X^TAX is spd.*

   Now consider the following A spd and blocking like:

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & w_i^T \\ w_1 & k_1 \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 & 0 \\ \frac{w_1}{\alpha_1} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & k_1 - w_1 \end{bmatrix} \begin{bmatrix} \alpha_1 & \frac{w_1^T}{\alpha_1} \\ 0 & I \end{bmatrix} \\ &= L_1 A_1 L_1^\top. \end{aligned}$$

   Use the last lemma to prove that A_1 spd implies $k_1 - \frac{w_1 w_1^T}{d}$ is spd

1. Show that if A is positive definite, then A^k is also positive definite for all $k \geq 1$.

Write A as a Jordan block

$$A = PJP^{-1}$$

$$J = \begin{bmatrix} \begin{bmatrix} \lambda_1 & 1 & & 0 \\ & \lambda_1 & 1 & \\ & & \ddots & \\ 0 & & & \lambda_1 & 1 \end{bmatrix} & & & 0 \\ & \begin{bmatrix} \lambda_2 & 1 & & \\ & \lambda_2 & 1 & \\ & & \ddots & \\ & & & \lambda_2 & 1 \end{bmatrix} & & \\ & & \ddots & \\ 0 & & & \begin{bmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & & \lambda_k & 1 \end{bmatrix} \end{bmatrix}$$

Too complicated.

Assume A is symmetric as well.

3. Let A be an $m \times n$ matrix with independent columns. If $A = QR$ and $A = \hat{Q}\hat{R}$ are QR-factorization of A, show that $Q = \hat{Q}$ and $R = \hat{R}$. Let's say $\hat{Q} : Q_1$ & $\hat{R} : R_1$

1. Write Q and Q_1 as:

$$Q : [c_1 \quad c_2 \quad \cdots \quad c_n] \qquad Q_1 : [d_1 \; d_2 \; \cdots \; d_n]$$

columns

Since Q & Q_1 are orthogonal

$$Q^T Q = I_n = Q_1^T Q_1$$

$\hookrightarrow$ id $n \times n$

$(R_1 = Q_1^T A = Q^T A = R)$

Then it's sufficient to show that $Q_1 = Q$.
Now since $Q_1^T Q_1 = I_n$ & since $QR = Q_1 R_1 \Rightarrow Q_1^T Q = R_1 R^{-1}$.
then since R_1 & R^{-1} are upper diagonal $\Rightarrow Q_1^T Q$ is upper diagonal & it has positive diagonal elements (because this is true for R and R_1).

$$(Q_1^T Q)_{ij} = t_{ij} = \begin{cases} t_{ii} > 0 \\ t_{ij} = 0 \quad \text{if } i > j \end{cases}$$

Notice that

$$Q_1^T = \begin{bmatrix} d_1^T \\ d_2^T \\ \vdots \\ d_n^T \end{bmatrix} \Rightarrow (Q_1^T Q)_{ij} = d_i^T c_j = d_i \bullet c_j$$

Then we have that $d_i^T c_j = t_{ij} \quad \forall\, i, j$.
But remember $Q = Q_1 (R_1 R^{-1})$

set of all linear combinations of the d_i's

linear combination of columns of Q_1 $\Rightarrow$ each column in Q (the c_i's) is a linear combination of the columns of Q_1.

$$[c_1 \; \cdots \; c_n] = Q_1 (R_1 R^{-1})$$

$\Rightarrow c_j \in \text{span}\{d_1, d_2, \ldots, d_n\}$
Because of how the QR factorization is done:

$$c_j = (d_1^T c_j) d_1 + (d_2^T c_j) d_2 + \ldots + (d_n^T c_j) d_n$$

"standard" a inner product $\langle a, b \rangle = \sum a_i b_i = a^T b$

$$= \langle d_1, c_j \rangle d_1 + \langle d_2, c_j \rangle d_2 + \ldots + \langle d_n, c_j \rangle d_n$$

$$= t_{1j} d_1 + t_{2j} d_2 + \ldots + t_{nj} d_n$$

$$= t_{1j} d_1 + t_{2j} d_2 + \ldots + t_{jj} d_j$$

Writing every c_j:

$$c_1 = t_{11} d_1$$

$$c_2 = t_{12} d_1 + t_{22} d_2$$

$$c_3 = t_{13} d_1 + t_{23} d_2 + t_{33} d_3$$

$$\vdots$$

set of linear equations!

Now:

$$\|c_1\| = \| t_{11} d_1 \| = |t_{11}| \, \|d_1\| = |t_{11}|$$

because the d_j's are the columns of an orthogonal matrix

$$\Rightarrow c_1 = d_1$$

Then for c_2

$$t_{12} = d_1^T c_2 = c_1^T c_2 = 0$$

Because the columns of an orthogonal matrix are orthogonal

$$\Rightarrow c_2 = t_{22} d_2 .$$

Again: $c_2 = d_2 \Rightarrow t_{13} = 0$ & $t_{23} = 0$

Continue with all the c_i's and get that

$$c_i = d_i \quad \forall i \Rightarrow Q_1 = Q$$

Because all the columns are the same.

$\blacksquare$