How to compute the QR factorization

$\underline{A} = \underline{\hat{Q}}\,\underline{\hat{R}}$ ← partial QR factorization.

$\underline{\hat{Q}}^T \underline{\hat{Q}} = \underline{\hat{Q}}\,\underline{\hat{Q}}^T = \underline{I} \Rightarrow \underline{\hat{Q}}^T = \underline{\hat{Q}}^{-1}$

Classical Gram-Schmidt Algorithm → projection based, not stable numerically

Turn to Modified Gram-Schmidt

Recall that projection can be written as a matrix-vector product

$\Rightarrow q_1 = \dfrac{\underline{P}_1 q_1}{\|\underline{P}_1 q_1\|}, \quad q_2 = \dfrac{\underline{P}_2 q_2}{\|\underline{P}_2 q_2\|}$, etc,

for some $\underline{P}_j$

Let $\underline{\hat{Q}}_{j-1}$ be the $m \times (j-1)$ matrix of the first $j-1$ columns of $\underline{\hat{Q}}$:

$\underline{\hat{Q}} = [q_1 \; q_2 \; q_3 \; \cdots \; q_n]$

$\underline{\hat{Q}}_{j-1} = [q_1, q_2 \; \cdots \; q_{j-1}]$

Then I can write $\underline{P}_j = \underline{I} - \underline{\hat{Q}}_{j-1}\underline{\hat{Q}}_{j-1}^T$

$\Rightarrow$ matrices of the form $\underline{I} - \underline{v}\,\underline{v}^T$ project onto the perpendicular space of $\underline{v}$.

Thus, $\underline{P}_j$ is nothing but the repeated perpendicular projections of each prior vector in $\underline{Q}$.

$$\underline{P}_j = \underline{P}_{\perp q_{j-1}} \underline{P}_{\perp q_{j-2}} \cdots \underline{P}_{\perp q_2} \underline{P}_{\perp q_1}$$

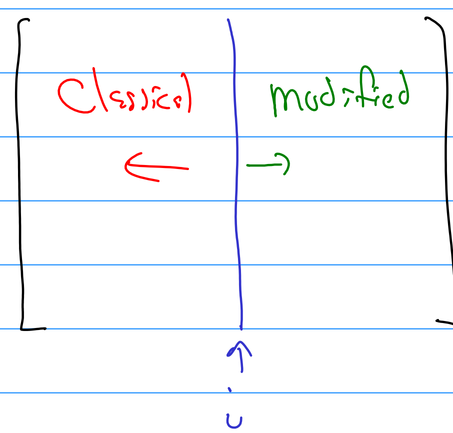
w/ $\underline{P}_1 = \underline{I}$

Each $\underline{P}_{\perp q_j}$ projects onto the space perpendicular to q_j.

Modified G-S uses these ideas to <u>reverse</u> the order of operation!

Algorithm: Modified G-S

for $i = 1:n$
  $\underline{v}_i = \underline{q}_i$
for $i = 1:n$
  $r_{ii} = \|\underline{v}_i\|$
  $\underline{q}_i = \underline{v}_i / r_{ii}$
  for $j = i+1:n$
    $r_{ij} = q_i^T \underline{v}_j$
    $\underline{v}_j = \underline{v}_j - r_{ij} q_i$



Op Count for Modified is the same as Classical G-S:

$$O(2mn^2)$$

Householder Triangularization

Look at G-S again.

In G-S, each operation to compute a column of $\hat{\underline{Q}}$ is an upper triangular matrix multiplication:

$$\underline{A}\,\underbrace{\underline{R}_1\,\underline{R}_2\cdots\underline{R}_n}_{\hat{\underline{R}}^{-1}} = \hat{\underline{Q}} \implies \underline{A} = \hat{\underline{Q}}\,\hat{\underline{R}}$$

This is called Triangular Orthogonalization or R given Q

You can do the reverse! repeated application of $\underline{Q}$ give $\hat{\underline{R}}$:

$$\underbrace{\underline{Q}_n\,\underline{Q}_{n-1}\cdots\underline{Q}_2\,\underline{Q}_1}_{\hat{\underline{Q}}^T}\,\underline{A} = \hat{\underline{R}} \implies \underline{A} = \hat{\underline{Q}}\,\hat{\underline{R}}$$

This is called orthogonal triangularization

Q gives R

Now we need to find the $\underline{Q}_k$.

The idea is to find a matrix $\underline{Q}_k$ that zeros out the values below a diagonal while preserving all prior zeros:

$$\underbrace{\begin{bmatrix} x & x & x \\ x & x & x \\ x & x & x \\ x & x & x \end{bmatrix}}_{\underline{A}} \xrightarrow{\underline{Q}_1} \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & x & x \end{bmatrix} \xrightarrow{\underline{Q}_2} \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & x \end{bmatrix} \xrightarrow{\underline{Q}_3} \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & 0 \end{bmatrix}$$

One more condition: $\underline{Q}_k$ must be unitary:

$$\underline{Q}_k^T \underline{Q}_k = \underline{Q}_k \underline{Q}_k^T = \underline{I}$$

Choose the following block matrix:

$$\underline{Q}_k = \begin{bmatrix} \underline{I} & \underline{0} \\ \underline{0} & \underline{F} \end{bmatrix}$$

$\underline{I} \in (k-1) \times (k-1)$ identity

$\underline{F} \in (m-k+1) \times (m-k+1)$ called a Householder reflector

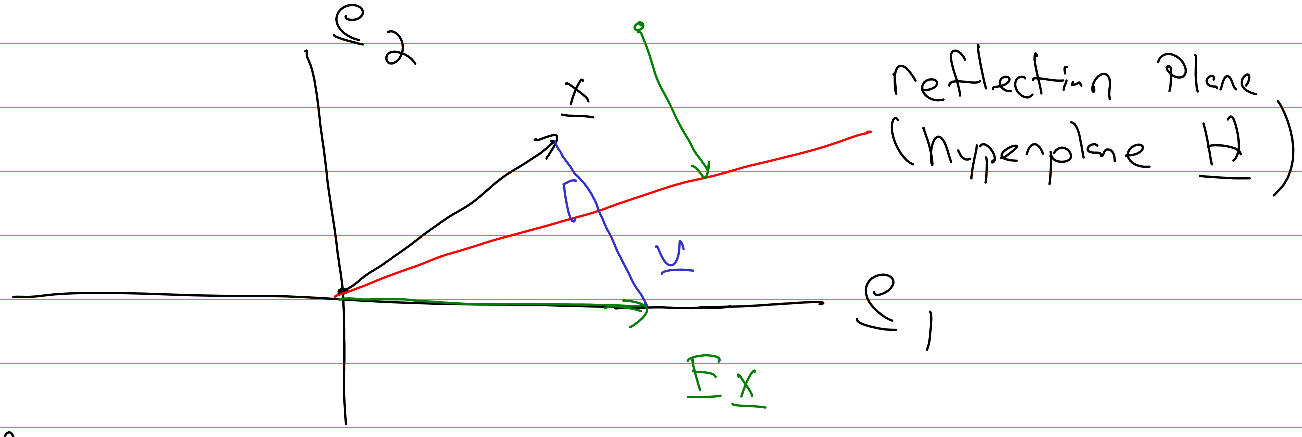
$\underline{F}$ is a specific type of operation:

It is defined as

$$\underline{x} = \begin{bmatrix} a \\ b \\ c \\ \vdots \\ \vdots \end{bmatrix} \xrightarrow{\underline{F}} \underline{F}\,\underline{x} = \begin{bmatrix} \|\underline{x}\|_2 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \|\underline{x}\|_2\, \underline{e}_1$$

$$\underline{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \end{bmatrix}$$

Look at it in 2D:



Hyperplane: A plane w/ a dimension one less than the embedding plane

(in 2D, H is 1D, in 3D H is 2D, etc.)

To determine this projection, look at the "error" vector between $\underline{F}\,\underline{x}$ & $\underline{x}$:

$$\underline{v} = \underline{F}\,\underline{x} - \underline{x} = \|\underline{x}\|_2\,\underline{e}_1 - \underline{x}$$

$\underline{v}$ is defined once $\underline{x}$ is.

The important part is that $\underline{v}$ is perpendicular to the hyperplane $\underline{H}$.

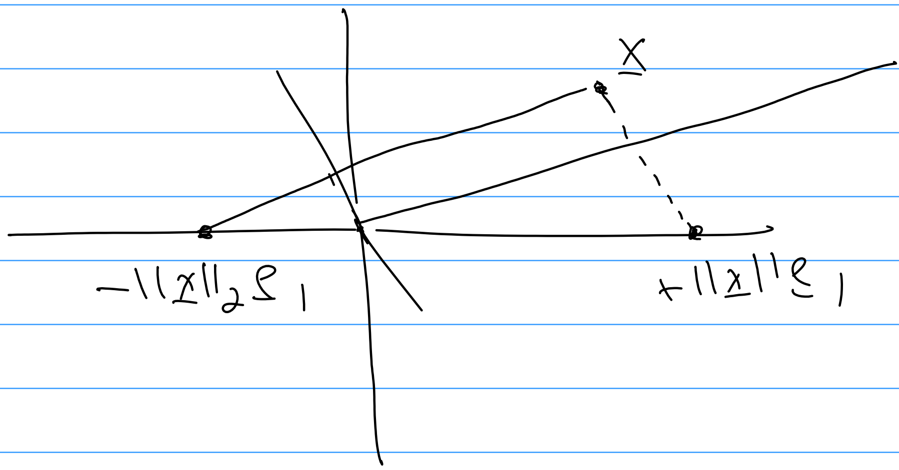
To develop $\underline{F}$, project a vector $\underline{y}$ onto $\underline{H}$:

$$\underline{P}\,\underline{y} = \left(\underline{I} - \frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}\right)\underline{y} = \underline{y} - \underline{v}\,\frac{(\underline{v}^T\underline{y})}{(\underline{v}^T\underline{v})}$$

But, we need to go twice as far.

$$\Rightarrow \underline{F}\,\underline{y} = \underbrace{\left(\underline{I} - 2\frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}\right)}_{\underline{F}}\underline{y} = \underline{y} - 2\underline{v}\left(\frac{\underline{v}^T\underline{y}}{\underline{v}^T\underline{v}}\right)$$

Note! Householder reflectors are not unique.



Mathematically, it doesn't matter.
Numerically, it does. We want large $\|\underline{v}\|$

$\Rightarrow$ Set $\underline{v} = -\text{sign}(x_1) \|\underline{x}\|_2 \underline{e}_1 - \underline{x}$

w/ $\text{sign}(x) = 1$ if $x = 0$

and x_1 = first component of $\underline{x}$.

After clearing the $\ominus$:

$\underline{v} = \text{sign}(x_1) \|\underline{x}\|_2 \underline{e}_1 + \underline{x}$

Algorithm: Householder QR

for $k = 1:n$

$\quad \underline{x} = \underline{A}(k:m, k)$

$\quad \underline{v}_k = \text{sign}(x_1)\,\|\underline{x}\|_2\,\underline{e}_1 + \underline{x}$

$\quad \underline{v}_k = \underline{v}_k / \|\underline{v}_k\|_2$

$\quad A(k:m, k:n) = A(k:m, k:n) - 2\underline{v}_k(\underline{v}_k^T \underline{A}(k:m, k:n))$

After this, $\underline{A}$ will be upper triangular.

Note: $\hat{\underline{Q}}$ is <u>never</u> computed.

To find $\hat{\underline{Q}}$ do the following:

Define the operation of $\hat{\underline{Q}}\,\underline{x}$

for $k = n:-1:1$

$\quad \underline{x}(k:m) = \underline{x}(k:m) - 2\underline{v}_k(\underline{v}_k^T \underline{x}(k:m))$

To find $\hat{\underline{Q}}$ apply this operation to the Identity:

$$\underline{Q}\,\underline{I} = \underline{Q} = (\underline{Q}\underline{e}_1, \underline{Q}\underline{e}_2 \cdots)$$

$$\underline{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \end{bmatrix} \quad \underline{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

Eigen Systems

To motivate this, look at the solution to

$$\frac{dy}{dt} = ay \quad , \quad a = \text{Constant}$$

Solution is $y(t) = Ce^{at}$

$$\frac{dy}{dt} = \frac{d}{dt}\left(Ce^{at}\right) = a\left(Ce^{at}\right) = a\,y(t)$$

What about a set of 2 ODEs?

$$\frac{dy_1}{dt} = ay_1 \qquad \frac{dy_2}{dt} = by_2$$

$$\Rightarrow y_1 = C_1 e^{at} \qquad y_2 = C_2 e^{bt}$$

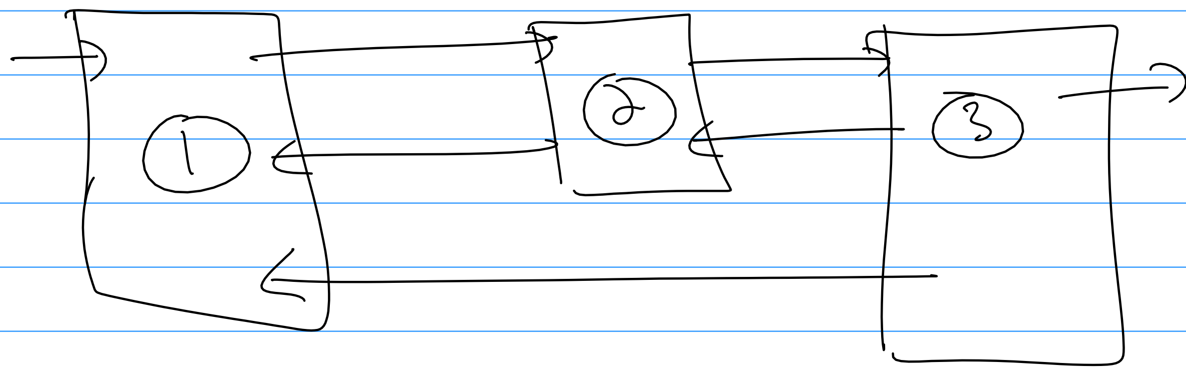
What if $\quad \dfrac{dy_1}{dt} = ay_2 \qquad \dfrac{dy_2}{dt} = by_1$

You need y_1 to solve for y_2, etc.

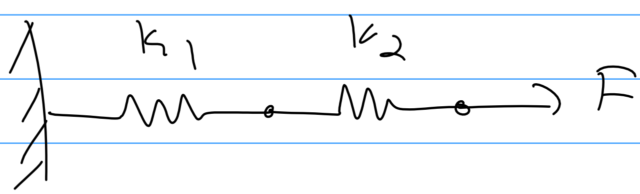
This is quite common

(1) A set of chemical reactions

② Flow between tanks:



③ Series of Springs:



Generalize these into the following:

$$\frac{dy_1}{dt} = a_{11} y_1 + a_{12} y_2$$

$$\frac{dy_2}{dt} = a_{21} y_1 + a_{22} y_2$$

Let $\underline{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ $\underline{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$$\Rightarrow \frac{d\underline{y}}{dt} = \begin{bmatrix} \partial y_1/\partial t \\ \partial y_2/\partial t \end{bmatrix} = \underline{A}\,\underline{y}$$

$\Rightarrow \frac{d\underline{y}}{dt} = \underline{A}\,\underline{y}$

Since $\frac{dy}{dt} = ay$ has a solution of $y(t) = Ce^{at}$

Make an ansatz that

$\underline{y}(t) = e^{\lambda t}\,\underline{x}$, for some constant vector $\underline{x}$,

Check:

$\frac{d\underline{y}}{dt} = \lambda e^{\lambda t}\,\underline{x}$

$\underline{A}\,\underline{y} = \underline{A}(e^{\lambda t}\underline{x}) = e^{\lambda t}\underline{A}\,\underline{x}$

$\frac{d\underline{y}}{dt} = \underline{A}\,\underline{y}$

$\lambda \cancel{e^{\lambda t}}\,\underline{x} = \cancel{e^{\lambda t}}\,\underline{A}\,\underline{x} \Rightarrow \underline{A}\,\underline{x} = \lambda\,\underline{x}$

If I can find $\underline{x}$ such that $\underline{A}\,\underline{x} = \lambda\,\underline{x}$,

then $\underline{y}(t) = e^{\lambda t}\,\underline{x}$ solves $\frac{d\underline{y}}{dt} = \underline{A}\,\underline{y}$

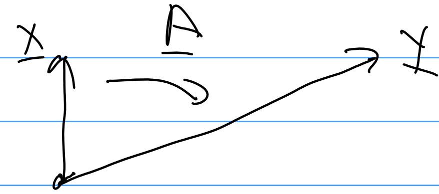
Systems of the form $\underline{A}\,\underline{x} = \lambda\,\underline{x}$ frequently appear,

Eigensystem of $\underline{A}$.

$\underline{x}$ = eigenvector $\qquad \lambda$ = eigenvalue.

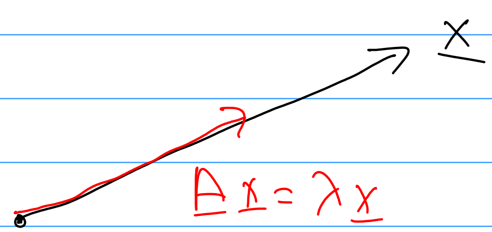
What does a matrix do in general?

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}}_{\underline{A}} \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\underline{x}} = \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{\underline{y}}$$



Matrices transform vectors.

Now look at $\underline{A}\,\underline{x} = \lambda\,\underline{x}$



$\underline{x}$ is the special set of vectors of $\underline{A}$ such that applying $\underline{A}$ to $\underline{x}$ does nothing but scale $\underline{x}$.

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}}_{\underline{A}} \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\underline{x}} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \underbrace{3}_{\lambda} \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\underline{x}}$$

Let $\underline{A} \in n \times n$ (eigen systems such as this are only valid for square matrices)

Given $\underline{A}$, find $\underline{x}$ & λ

$$\underline{A}\,\underline{x} = \lambda \underline{x}$$

$$\underline{A}\,\underline{x} - \lambda \underline{x} = \underline{0}$$

$$(\underline{A} - \lambda \underline{I})\underline{x} = \underline{0}$$

$\Rightarrow$ $\underline{x}$ must be in the null space of $\underline{A} - \lambda \underline{I}$.

We do <u>not</u> want the trivial solution $\underline{x} = \underline{0}$.

$\Rightarrow$ $\underline{A} - \lambda \underline{I}$ can <u>not</u> be full rank.

$\Rightarrow$ $(\underline{A} - \lambda \underline{I})^{-1}$ does <u>not</u> exist

$\Rightarrow$ $\det(\underline{A} - \underset{\uparrow \text{ unknown}}{\lambda} \underline{I}) = 0 \leftarrow$

$\det(\underline{A} - \lambda \underline{I})$ = Characteristic polynomial of $\underline{A}$.

Eigensystem procedure: Given $\underline{A} \in n \times n$

(1) Solve for all λ's such that

$$\det(\underline{A} - \lambda \underline{I}) = 0$$

(2) for each λ_i find $\underline{x}_i$ such that

$$\underline{A}\,\underline{x}_i = \lambda_i \underline{x}_i$$

or

$$(\underline{A} - \lambda_i \underline{I})\underline{x}_i = 0$$

ex.) Let $\underline{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

$$\underline{A} - \lambda \underline{I} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix}$$

$\det(\underline{A}) = |\underline{A}|$

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 4 = 0$$

$(1-\lambda)^2 = 4$, $1-\lambda = \pm 2$

$\Rightarrow \lambda_1 = -1$ & $\lambda_2 = 3$

for $\lambda_1 = -1$, "Solve" $(\underline{A} - \lambda \underline{I})\underline{x} = 0$

$$\begin{bmatrix} 1-\lambda_1 & 2 \\ 2 & 1-\lambda_1 \end{bmatrix} \underline{x}_1 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \underline{x}_1 = \underline{0}$$

$$\text{rref}(\underline{A} - \lambda_1 \underline{I}) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \underline{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

for $\lambda_2 = 3$: $\begin{bmatrix} 1-\lambda_2 & 2 \\ 2 & 1-\lambda_2 \end{bmatrix} \underline{x}_2 = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \underline{x}_2 = \underline{0}$

$$\Rightarrow \underline{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

for $\underline{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, there are 2 eigenvalues:

$\lambda_1 = -1$, $\underline{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $\qquad \lambda_2 = 3$, $\underline{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = (-1)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = (3)\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$