**Questions**

Problems

1. A is ill-conditioned → residuals are not a good measure of fit
2. We should be looking at the precision of each fit
   1. Ie: how many decimal points are the same

**Questions that we should answer in our presentation:**

Why are the K(A)s different for each A matrix?

* Def Condition number
  + determines the loss in precision due to roundoff errors in Gaussian elimination and can be used to estimate the accuracy of results obtained from matrix inversion and linear equation solution
  + measure of how ill-conditioned A is and can be found using A and A^−1
  + Measure how invertible A is
    - Small K(A) → more invertible
  + Measures how much the output value of the function can change for a small change in the input argument
    - We change our data to include some noise, see a drastic change in condition number (wait why?? We only changed the vector b, which does not change the condition number of A)
    - what did we see a drastic change in?
  + Matrix A is ill-conditioned if it is invertible but can become non-invertible (singular) if some of its entries are changed ever so slightly
    - Small change in inputs → large change in output
    - ie: correct solution/answer becomes hard to find
  + Well-conditioned matrices have condition numbers close to 1

What is the intent of including noise into our data?

* Want to see if a small change in our dataset will affect our fit of the model
* Essentially empirically testing the statement that our condition number makes
  + If our condition number is high, does a small change in the data (vector b) result in a big difference in our fitted model/accuracy of our fitted model?
* Apparently, if our matrix is ill-conditioned → **hard to use residuals as a measure of accuracy**

**Conclusions for us to make:** (from what I can tell so far) - Gianni

* Yes we have some high condition numbers for polynomial models
* “If we define relative residual as ∥r∥∥b∥, we can see that small relative residual implies small relative error in approximate solution only if A is well-conditioned (cond(A) is small).”
  + Our condition number for

Perturbations

* Both prove instability
  + Perturbation in b
  + Perturbation in A

Compare the accuracy of fit between the normal equation ( vs QR decomposition vs SVD

* Fit is same

What are the assumptions you made before running the data and what did you expect to see?

Condition number vs accuracy

**Questions that we could be asked:**

Why did we chose to compare the performance of each method using instead of ?

* Both should solve for the same thing
* Wanted to see the fit of x, don’t want rounding errors of QR

How would you set up your A matrix if you wanted to add a second predictor variable?

* Add the column into our matrix A

Why are the estimates of the vector x the same for CGS, MGS, Householder, and SVD?

Why did you decide to approach Least Squares with QR Factorization and SVD? What are the advantages and disadvantages of each method?

What are some disadvantages of your model that you’d like to note?

* Overfitting
  + Cubic model outperforms linear model bc it can fit the values on the y-axis

What are the assumptions you made before running the data and what did you expect to see?

If you could redo this project again, what changes would you make?

Qs about the presentation:

* How in depth should our explanation of each method be? Are we essentially reteaching CGS, MGS, and Householder?

**Perturbation and Condition Numbers**

What are Condition Numbers?

* determines the loss in precision due to roundoff errors in Gaussian elimination and can be used to estimate the accuracy of results obtained from matrix inversion and linear equation solution
* measure of how ill-conditioned A is and can be found using A and A^−1
* Measure how invertible A is
  + Small K(A) → more invertible
* Measures how much the output value of the function can change for a small change in the input argument

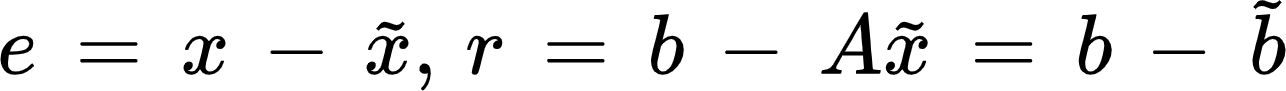
When can we not use residuals as measure of goodness of fit and why?

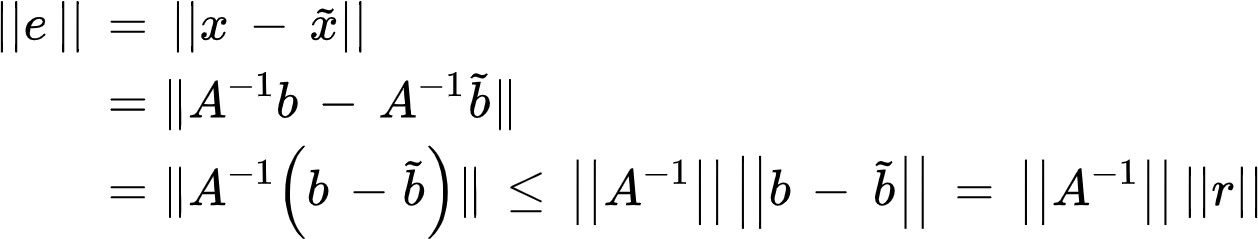
* When A is ill-conditioned (ie: K(A) is large), the residual can be small even though the answer clearly wrong
* To prove that A is ill-conditioned, we add more noise to the b vector
  + If the residual changes drastically, we can further confirm that A is ill-conditioned

Derivation of why the residuals are not accurate indicators of the error

* In this case, we assume that A is square (or else the K(A) will be infinite)

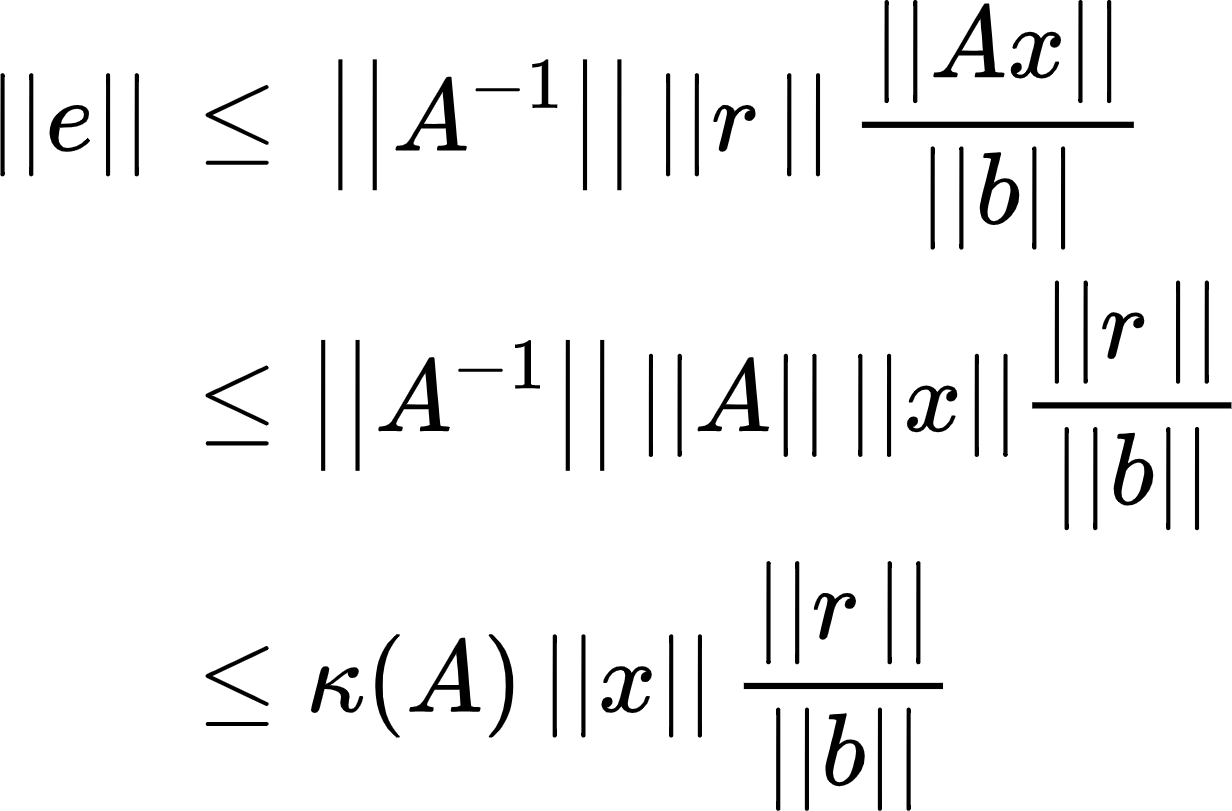
x\_tilde = estimated x



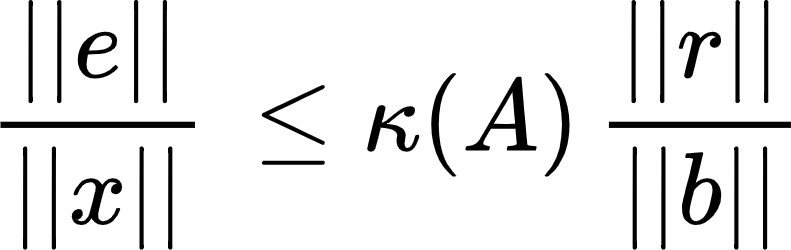


So,

However, when we consider the relative error



Dividing both sides with the norm of x, we get:



* Relative error is dependent on the condition number
* If K(A) is large (ie: ill-conditioned), then the small perturbations in r (and b) will lead to large relative error
  + Hence, the residual is an unreliable indicator of goodness of fit

Note: Since our A matrix is not square → cannot find the inverse → must be ill-conditioned