Module 7

Modern Navigation Systems

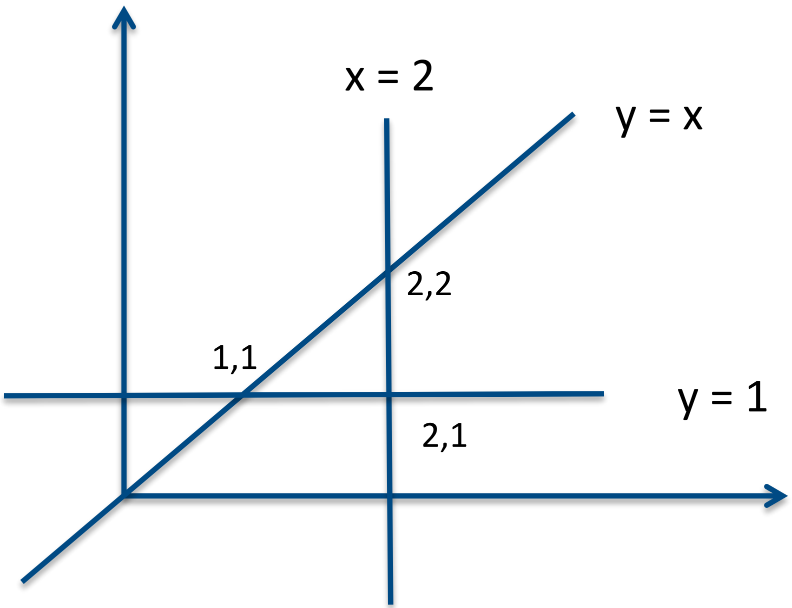
Lines of position and weighted least squares

Module 7C
Weighted Least Squares

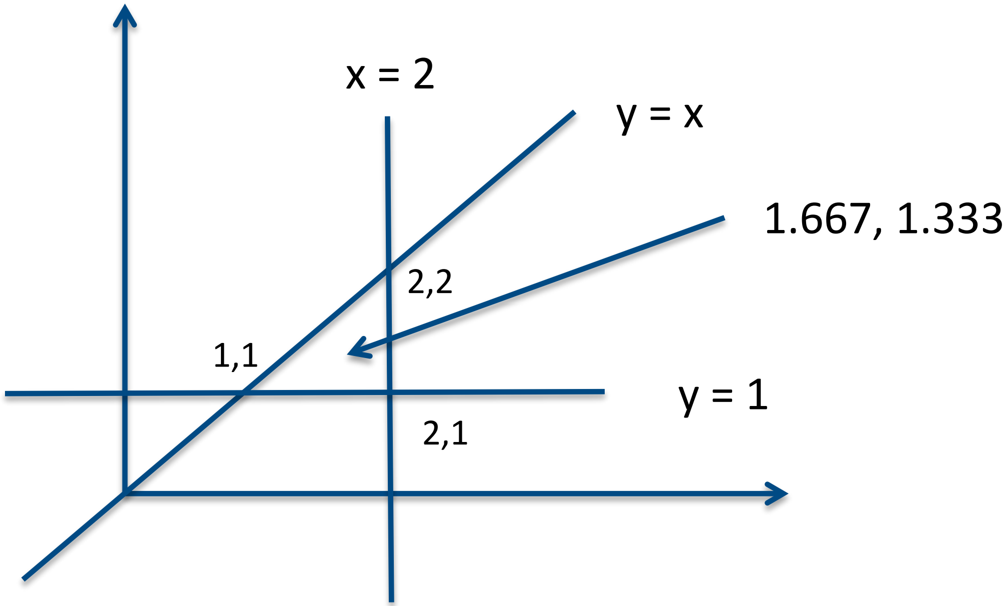
Summary of Module 7

- Students will combine the perturbation analysis and position determination skills from previous modules to define lines of position. The goal is to use the techniques of linear algebra to compute the best estimate of the point in three dimensional space at which the lines intersect. This will yield a “position fix”. (7A)
- The techniques are then extended to the case where the problem is over-specified. That is, m measurements are made with respect to the solution of an n-dimensional set of equations, where m > n. This leads to the least-squares solution of the problem. (7B)
- **The concept of measurement uncertainty is introduced, leading to the technique of weighted least squares. This leads to the concept of the Kalman filter, which is introduced briefly. (7C)**

Solving 3 equations in 2 unknowns that don't intersect at the same point:



Solve this using matrices

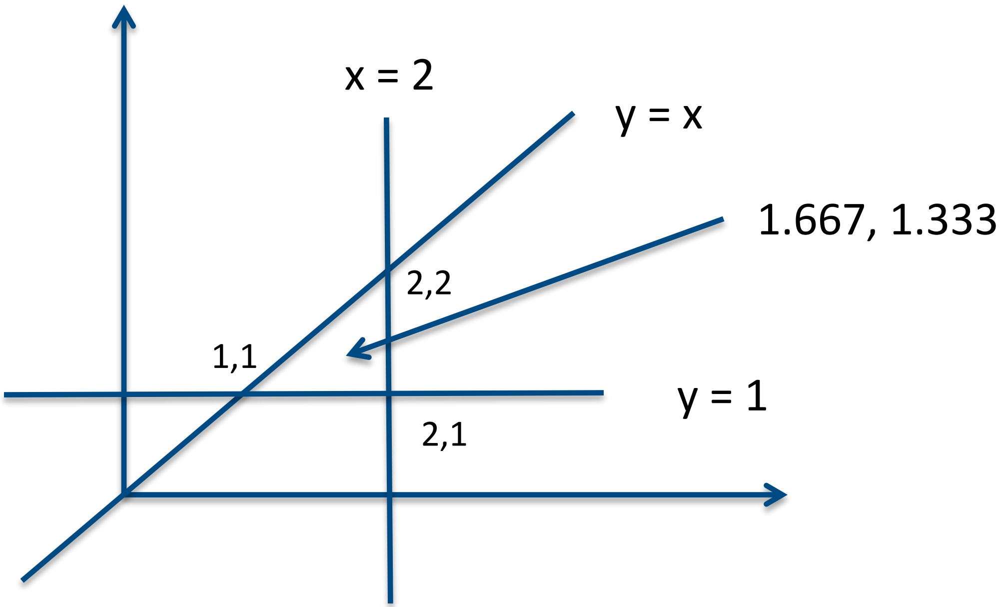


Define the lines using $Ax = B$, where A is a matrix and x and B are vectors.

A will have 3 rows and two columns, making the system of equations unsolvable.

Solve the system $A^TAx = A^TB$ instead. This will yield the least squares solution, which minimizes the sum of the squares of the minimum distances from the solution point to each of the original lines.

Solve this using matrices, cont'd



Three equations in 2 unknowns:

$0x + y = 1$

$x + 0y = 2$

$x - y = 0$

$$A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \qquad B := \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

$$A^T = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

$$x := \left(A^T \cdot A\right)^{-1} \cdot A^T \cdot B$$

$$x = \begin{pmatrix} 1.667 \\ 1.333 \end{pmatrix}$$

But what if the lines of position are affected by measurement error…

- Each of the lines represents the results of a measurement
  - Distance to a GPS satellite
  - Angle above the horizon of a star
  - Bearing to a terrestrial radio beacon or visual landmark
  - Compass measurement
  - etc.

What if these measurements differ in precision and accuracy?

- It is possible to weight the different measurements by inserting a new matrix:
  - $Ax = B$ becomes $A^T Ax = A^T B$ becomes
    - $A^T WAx = A^T WB$, which is solved using
      - $x = (A^T WA)^{-1} A^T WB$
      - Where W is a square matrix that has equal numbers of rows and columns

For example, let W be the identity matrix

Three equations in 2 unknowns:

$0x + y = 1$

$x + 0y = 2$

$x - y = 0$

$$A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \qquad B := \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

$$A^T = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \qquad W := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A^T \cdot A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad A^T \cdot W \cdot A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$x := \left(A^T \cdot W \cdot A\right)^{-1} \cdot A^T \cdot W \cdot B$$

$$x = \begin{pmatrix} 1.667 \\ 1.333 \end{pmatrix}$$

Note that this choice of W has no impact on the solution to the problem!

Now, set one row of W to all zeros

Three equations in 2 unknowns:

$0x + y = 1$

$x + 0y = 2$

$x - y = 0$

$$A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \qquad B := \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

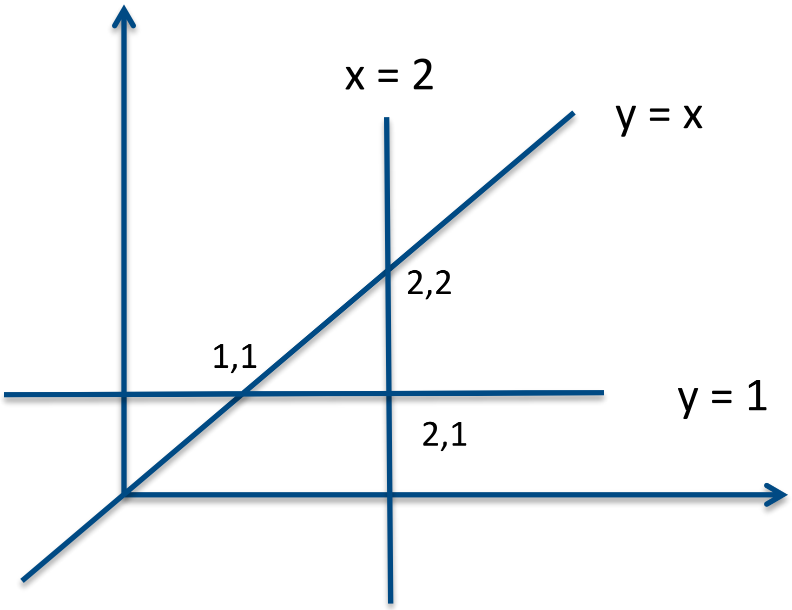
$$A^T = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \qquad W := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A^T \cdot A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad A^T \cdot W \cdot A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

$$x := \left(A^T \cdot W \cdot A\right)^{-1} \cdot A^T \cdot W \cdot B$$

$$x = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

Note that this choice of W turns off the line y = 1.



Now, set a different row of W to all zeros

Three equations in 2 unknowns:

$0x + y = 1$

$x + 0y = 2$

$x - y = 0$

$$A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \qquad B := \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

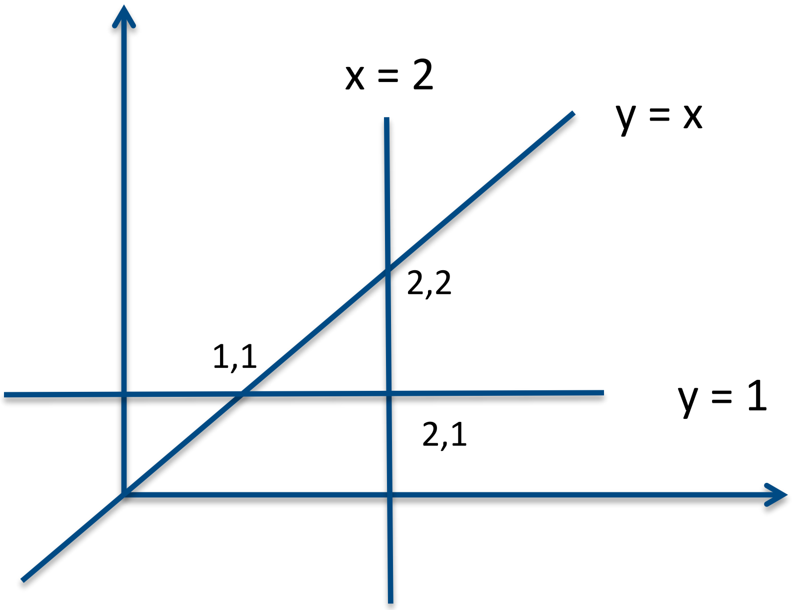
$$A^T = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \qquad W := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$A^T \cdot A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad A^T \cdot W \cdot A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$x := \left(A^T \cdot W \cdot A\right)^{-1} \cdot A^T \cdot W \cdot B$$

$$x = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Note that this choice of W turns off the line y = x.



Now, set off-diagonal elements of W to non-zero values, and change the diagonal values to numbers other than 1:

Three equations in 2 unknowns:

$0x + y = 1$

$x + 0y = 2$

$x - y = 0$

$$A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \qquad B := \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

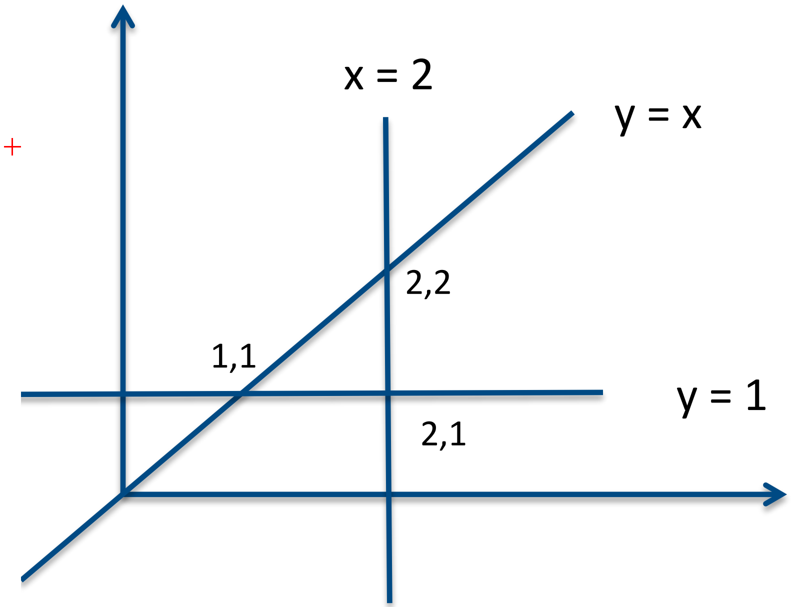
$$A^T = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \qquad W := \begin{pmatrix} 1 & .2 & 0 \\ 0 & 1 & 0 \\ 0.6 & 0 & .5 \end{pmatrix}$$

$$A^T \cdot A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad A^T \cdot W \cdot A = \begin{pmatrix} 1.5 & 0.1 \\ -0.3 & 0.9 \end{pmatrix}$$

$$x := \left(A^T \cdot W \cdot A\right)^{-1} \cdot A^T \cdot W \cdot B$$

$$x = \begin{pmatrix} 1.638 \\ 1.435 \end{pmatrix}$$

Note that this choice of W still provides a solution that lies within the triangle below.



What is the best choice of W?

- Several texts address this problem. One of the best is *Intro to Applied Mathematics*, by Strang.
- The primary text for this course, by Kayton & Fried, addresses this in many places, starting with section 2.8, Navigation Errors. (Consult the index of the text, including sections on Kalman filters, for additional relevant sections of the text.)
  - Note the introduction of the term $\sigma_x\sigma_y$
    - σ is the standard deviation
    - σ^2 is the variance
    - $\sigma_x\sigma_y$ is the covariance of the variables *x* and *y*

What if measurements are not orthogonal (that is, they are not independent of each other)?

- The lines y = 1 and x = 2 are at right angles
  - y = 1 provides no information about x
  - x = 2 provides no information about y
  - Hence, the covariance between noisy measurements of each line should be 0, if different instruments or independent observations are made.
- However, an error in the measurement of the line y = x will affect estimates of both x and y!
  - Hence, $\sigma_x\sigma_y$ for this case will be non-zero, and a corresponding non-zero entry in a matrix of covariance values should be made

Example of a 2 x 2 covariance matrix

$$C = \begin{matrix} \sigma_x \sigma_x & \sigma_x \sigma_y \\ \sigma_y \sigma_x & \sigma_y \sigma_y \end{matrix}$$

It turns out that the best choice of W is that it be the inverse of the covariance matrix, namely C^{-1}

- This choice represents the optimal solution (least weighted-mean-squared error) of a linear system
- It is readily adapted to systems in which multiple measurements are averaged
  - but for which various measurements are made with different values of precision and/or accuracy
  - or the various measurements are only partially independent
    - For example, a fever and an elevated white count due to an illness are often correlated
    - Witnesses to a crime who know each other are liable to affected each other’s recollection of the incident

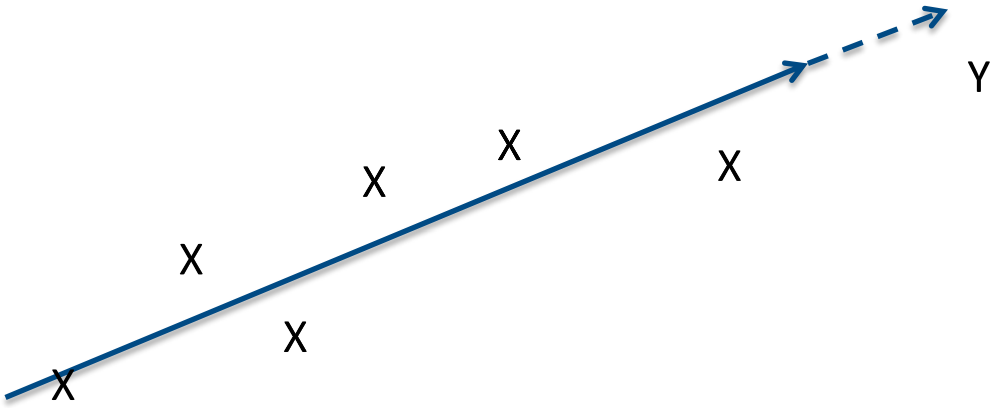
Relevance of covariance to navigation

- The total measurement error for a navigation fix is related to the square root of the sums of the variances of the individual measurements that contribute to computation of the fix, as shown in section 2.8 of the text.
- Measurements of lesser accuracy (i.e., measurements that have a higher variance) should have a reduced impact on the final estimate of one’s position, velocity, etc.
- The use of weighted least squares addresses this in an optimal fashion.

The Kalman Filter

- Suppose one wishes to average several noisy measurements to reduce uncertainty, but the parameter being measured is changing as a function of time
- For example, an aircraft is flying while being observed by a radar, but the radar data is noisy
- By fitting the radar data to a “motion model”, it is possible to average the errors in the radar estimates of the aircraft’s position, even though the aircraft has moved between each measurement
- A graphical example captures this well…

Prediction versus measurement…



The x's represent noisy radar measurements of an aircraft's position at different points in time.

The y represents a prediction, based on the least-squares averaging of the x's, of where the aircraft will be found at the next radar measurement at the next time increment. In order to average the x's, the data must be fit to a motion model, namely an estimate of the aircraft's speed and direction of travel.

As measurements are added, each new y has a reduced impact on the relative weighting of the tracking estimate predicted by the averaging of a large number of previous observations (i.e., the x's).

A Kalman filter is the optimal algorithm for combining averages of past measurements with new measurements, while using weighted least squares, and while continually updating the covariance estimates based on the fluctuations in this averaged data. It's success depends on the quality of data and the accuracy of the motion model.

The Kalman filter

- The Kalman filter algorithm is quite complicated, and is the subject of a separate,700 level course in the EP program.
- A key feature of the algorithm is its use of recursive techniques to minimize the computational burden.
- To explore this, the simplest case is given in section 3.9.1 of the primary text. There is no motion, and only data from a single sensor is processed. This simple example is extended systematically in the discussion that follows in chapter 3, with a block diagram of a complete Kalman filter given in Figure 3.8.

Assignment 7.3

7.3.1 Using random numbers, and setting the deterministic error to zero, simulate equations 3.1 through 3.6 of the text.

7.3.2 Extend this to simulation of equations 3.7 to 3.11. Make the simulated errors in one set of measurements significantly larger than those in the other, thus illustrating that inclusion of measurements with low weights is in principle helpful, but in practice is not always worth doing.

7.3.3 Comment on the relationship of the weighted least squares solution of an over-specified problem to the signal averaging technique developed in problems 1 and 2, and in chapter 3 of the primary text.

End of Mod 7C