UNIVERSITY OF CALIFORNIA, BERKELEY

Chapter 6: Linear Equations

Oscar Ortega

July 16, 2021

1 6.1: THE SET OF SOLUTIONS OF LINEAR EQUATIONS

solution set : $S = \{x \in \mathbb{R}^n : Ax = y\}$

$$\text{Let } A = \begin{vmatrix} a_1 & a_2 & \dots & a_n \end{vmatrix}$$
$$\rightarrow Ax = \begin{vmatrix} x_1 a_1 + \dots + x_n a_n \end{vmatrix}$$

Recall:

$$\text{Range}(A): \text{ space spanned by the columns of a}$$

Rank Test:

$$rank(Ay) = rank(A) \rightarrow y \in R(A)$$

Formally: The linear equation Ax = y admits a solution iff $rank(Ay) = rank(A)$. When this existence condition is satisfied, the set of all solutions is the affine set:

$$S = \{x = \bar{x} + z : z \in \mathcal{N}(A)\}$$

2 6.2: UNDERDETERMINED, OVERDETERMINED, AND SQUARE SYSTEMS

Theorom:

$$A \in \mathbb{R}^{(m,n)} \text{is full column rank (i.e rank(A) = n) iff and only if } A^T A \text{is invertible}$$

$$A \in \mathbb{R}^{(m,n)} \text{is full row rank(i.e, rank(A) = m) iff and only if } AA^T \text{is invertible.}$$

Overdetermined system: Can think of a skinny matrix, m > n.
Underdetermined system: can think of fat matrix
Square systems:
If full column rank we can define the inverse of A as the matrix B s.t $AB = BA = I$.

3 APPROXIMATE SOLUTIONS: LEAST SQUARES

In systems where the solution set is empty, it makes sense to determine an approximate solution:

We consider the residual vector $r = Ax - y$ as the sort of error between our data vector and the multiplication of our data matrix A with our solution vector x. We want to minimize r: This implies the formal definition:

$$min_x \|Ax - y\|_2$$

Recall:
x^2 is monotonically increasing on the set of positive numbers, this implies we can minimize the square of this and achieve the same answer.

$$\|Ax - y\|_2^2 = \sum_i (a_i x - y_i)^2$$

this is were we get the name least squares:
Can also interpret a finding a point $\bar{y} \in \mathscr{R}(A)$ that is closest to y. This can be thought of as the orthogonal projection of y onto $\mathscr{R}(A)$
can perform calculus and optimize over x to yield the least squares solution to X.

$$x* = (A^T A)^{-1} A^T y$$

4 THE UNDERDETERMINED CASE: MINIMUM-NORM SOLUTION

: Recall: solution set of : $= \{x : x = \bar{x} + z, z \in \mathscr{N}(A)\}$ When it is the case we have an infinite number of solutions: which one do we choose? The smallest (simplest one). In other words: we want the solution with z = 0.

$$x * \text{must be orthogonal} to \mathscr{N}(A)$$

$$\rightarrow x* \in \mathscr{R}(A^T)$$

$$\rightarrow x* = A^T \gamma$$

Because we need x* to solve the system of equations: we need $Ax = y \rightarrow A(A^T \gamma) = y$

$$\rightarrow \gamma = (AA^T)^{-1} y$$

$$\rightarrow x* = A^T (AA^T)^{-1} y$$

5 VARIANTS OF THE LEAST-SQUARES PROBLEM

Linear equality-constrained LS: A generalization of the basic LS problem(6.5) allows for the addition of linear equality constraints on the x variable, resulting in the constrained problem:

$$min_x \|Ax - y\|_2^2 \text{s.t} : Cx = d$$

Will learn techniques for this later in the class.
6.7.2 (Weighted LS):

$$\text{W} = diag(w_1, ..., w_m)$$

new problem: $min_x \|W(Ax - y)\|_2$ This is also known as Tikhonov regularization: Based on the bayesian interpretation on the apriori belief that certain values of x are more likely than others.
6.7.3(l_2 regularized LS) new problem:

$$min_x \|Ax - y\|_2^2 + \lambda \|x\|_2^2$$

Idea: We want to find a solution that is close to the vector y without making the norm of x to large.