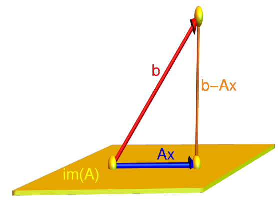
LINEAR ALGEBRA

MATH 21B



Data fitting

15.1. Data fitting problems lead to systems of equations $A\vec{x} = \vec{b}$ which do not need to have solutions. They could would tell for example that the data points are on a specific line. If the points are not part of a line there is no solution. But we can look for the point $A\vec{x} \in \text{im}(A)$ that is closest to $\vec{b}$. In other words, we like to find the projection of $\vec{b}$ onto the image of A.

15.2. Let $A\vec{x}$ be on $V = \text{im}(A)$, closest to $\vec{b}$. Then $\vec{b} - A\vec{x}$ is in $V^\perp = \text{im}(A)^\perp = \ker(A^T)$. This means $A^T(\vec{b} - A\vec{x}) = 0$. This leads to to $A^T\vec{b} = A^T A\vec{x}$. We can now solve for $\vec{x}$ and get the explicit **least square solution**

$$\vec{x} = (A^T A)^{-1} A^T \vec{b}\,.$$

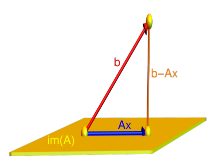


Figure 1. $A\vec{x} - \vec{b}$ perpendicular to $\text{im}(A)$ means it is in $\ker(A^T)$ leading to the least square solution.

15.3. Problem:
The first 7 prime numbers 2, 3, 5, 7, 11, 13 define the data points $(1,2), (2,3), (3,5), (4,7), (5,11), (6,13)$ in the plane. Our task is to find the best line $y = ax + b$ which fits these data.

Solution: Lets write down the equations for the unknown quantities a, b. We have

$$\begin{aligned} a \cdot 1 + b &= 2 \\ a \cdot 2 + b &= 3 \\ a \cdot 3 + b &= 5 \\ a \cdot 4 + b &= 7 \\ a \cdot 5 + b &= 11 \\ a \cdot 6 + b &= 13 \end{aligned}$$

The matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \end{bmatrix}$ and $\vec{b} = [2, 3, 5, 7, 11, 13]$. We can use the data fitting formula to get $[306, -76]/168$.

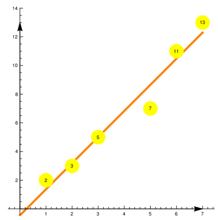


FIGURE 2. We want to find the best linear function fitting the first few points of the prime sequence $(k, p(k))$, where $p(k)$ is the k'th prime.

15.4. Problem. Find a quadratic polynomial $p(x) = ux^2 + vx + w$ which best fits the four data points $(-1, 8), (0, 8), (1, 4), (2, 16)$.
Solution. We write down the equations $ux^2 + vx + w = y$, for every data point (x, y). This gives us a system of four equations $Ax = b$ with $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$ $b = \begin{bmatrix} 8 \\ 8 \\ 4 \\ 16 \end{bmatrix}$.
$A^T A = \begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix}$. The solution is $x^* = (A^T A)^{-1} A^T b = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}^T$.

15.5. This gives the projection formula $\boxed{P\vec{v} = A(A^T A)^{-1} A^T \vec{b}}$ onto $\text{im}(A)$.

15.6. In the formula for the least square solution we need to worry about the fact that $A^T A$ is invertible. But this is equivalent to A has a trivial kernel and is usually the case. here is a situation, where A has a kernel:

15.7. Problem: Analyze the best linear fit $f(x) = ax + b$ for the three data points $(1, 1), (1, 2), (1, 3)$.

Solution: Write down the system $a + b = 1, a + b = 2, a + b = 3$. The matrix A has a kernel and $A^T A$ is not invertible. What happens is that there are infinitely many least square solutions to this problem.

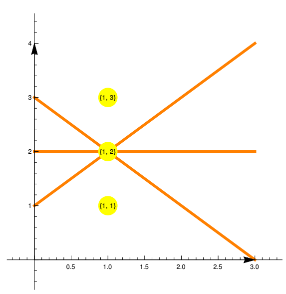


FIGURE 3. All non-vertical lines through $(1, 2)$ are least square solutions of the regression data fitting problem.