

Exercises

§5.2: Least Squares in Several Variables

Python: There is an optional starter script that supports Python-based problem(s) on this sheet. Use it as scaffolding: read the comments, complete the lines marked **TODO**, and compare the output with the mathematics requested in the problem(s).

		Observation	1	2	3	
Problem 1	Consider the data	x_1	1	2	3	For the model $\hat{y} = w_0 + w_1x_1 + w_2x_2$, write the design matrix X , response vector $\mathbf{y}$, and parameter vector $\mathbf{w}$. State the dimensions of each.
		x_2	0	1	1	
		y	2	5	7	

Problem 2	For $X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$, compute $X^T X$ and $X^T \mathbf{y}$. Solve the normal equations for $\mathbf{w}$.
------------------	---

Problem 3	For the two-predictor example from class, $\mathbf{w}^* = \begin{bmatrix} 3/4 \\ 5/2 \\ 7/2 \end{bmatrix}$. Compute $X\mathbf{w}^*$ and the residual vector $\mathbf{r}$. Verify directly that $X^T \mathbf{r} = \mathbf{0}$.
------------------	---

Problem 4 Let $X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & 3 & 6 \end{bmatrix}$. Explain why the columns of X are linearly dependent. What does this imply about $X^T X$ and uniqueness of the least-squares coefficients?

Problem 5 Suppose X has full column rank. Show that for every nonzero vector $\mathbf{v}$, $\mathbf{v}^T X^T X \mathbf{v} > 0$. Conclude that $X^T X$ is positive definite.

Problem 6 Use Python and np to fit the data

```
x = np.array([[1., 1., 0.],  
              [1., 2., 1.],  
              [1., 3., 0.],  
              [1., 4., 1.]])
```

```
y = np.array([3., 6., 7., 10.])
```

Report the coefficient vector. Then compute X and explain the result.