Title: Extra Homework, Week 12
Author: Cretu Cristian, 913

$A \in \mathbb{R}^{m\times n}$ is a full rank matrix with $m < n$.

1

We prove that AA^T is positive semi-definite using the definition: $x^T AA^T x \geq 0.$, where $x \in \mathbb{R}^m$.

A has full rank $\Rightarrow Ax \neq 0$ for all $x \neq 0 \Rightarrow x^T AA^T x = (Ax)^T(Ax) = \|Ax\|^2 \geq 0 \Rightarrow x^T AA^T x \geq 0$

We prove that AA^T is positive definite, $x^T AA^T x > 0$. using the same argument as before for $x \neq 0$, $Ax \neq 0$ since A has full rank $\Rightarrow x^T AA^T x = \|Ax\|^2 > 0 \Rightarrow AA^T$ is positive definite.

A positive definite matrix is invertible. Since AA^T is positive definite $\Rightarrow AA^T$ is also invertible.

2

$$\min_{x\in\mathbb{R}^n} \|x\|^2 \quad \text{subject to} \quad Ax = b.$$

$$f(x) = \|x\|^2 = x^T x, g(x) = Ax - b$$
$$L(x, \lambda) = x^T x - \lambda^T(Ax - b)$$

$$\frac{\partial L}{\partial x} = 2x - A^T\lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = Ax - b = 0$$

$$x = \frac{1}{2}A^T\lambda$$
$$Ax = b$$

$$\Rightarrow \frac{1}{2}AA^T\lambda = b$$

Since AA^T is invertible (from part 1), we can solve for λ, without the $\frac{1}{2}$ factor:

$$\lambda = (AA^T)^{-1}b.$$

Substituting back for x $\Rightarrow x^* = A^T\lambda = A^T(AA^T)^{-1}b$

3

Since x and x^* are both solutions for $Ax = b \Rightarrow Ax = Ax^* = b$. We substitute $\Rightarrow A(x - x^*) = 0$. THis means that $(x - x^*)^T x^* = 0$.

Also, since x^* minimizes $\|x\|^2$ and is a unique solution for the min. norm this means that $\|x^*\| \leq \|x\|$.

4

$$\min_{x \in \mathbb{R}^n} \left(\|Ax - b\|^2 + \alpha \|x\|^2 \right)$$

We calculate the derivative with respect to x, and compare it with 0 to find the minimum of the function:

$$f(x) = \|Ax - b\|^2 + \alpha\|x\|^2 = (Ax - b)^T(Ax - b) + \alpha x^T x.$$

$$\frac{\partial f}{\partial x} 2A^T(Ax - b) + 2\alpha x = 0.$$

$$A^T Ax + \alpha x = A^T b.$$

$$(A^T A + \alpha I)x = A^T b.$$

$$\Rightarrow x_\alpha = (A^T A + \alpha I)^{-1} A^T b$$

5

$A^T A + \alpha I$ is positive definite if:

$$x^T(A^T A + \alpha I)x > 0.$$

$$x^T(A^T A + \alpha I)x = x^T A^T Ax + \alpha x^T x = (Ax)^T(Ax) + \alpha x^T x = \|Ax\|^2 + \alpha\|x\|^2.$$

$\|Ax\|^2 \geq 0$ and $\alpha\|x\|^2 > 0 \Rightarrow A^T A + \alpha I$ is positive definite. This also results that $A^T A + \alpha I$ is invertible.

$$x_\alpha = (A^T A + \alpha I)^{-1} A^T b.$$

We apply the limit to both sides: as α approaches 0, αI approaches 0 as well.

$$\lim_{\alpha \to 0} x_\alpha = (A^T A)^{-1} A^T b = x^*.$$

6

[39]:
```
from sklearn.linear_model import Ridge
from sklearn.datasets import load_breast_cancer
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
```

[40]:
```
# load dataset
x, y = load_breast_cancer(return_X_y=True)
```

[69]:
```
# split 80/20
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2,␣
↪random_state=17)
```

[70]:
```
# make alpha 1.0
ridge_regression = Ridge(alpha=1.0)
```

[71]:
```
# fit the model
ridge_regression.fit(x_train, y_train)
```

[71]:
```
Ridge()
```

[72]:
```
y_pred = ridge_regression.predict(x_test)
```

[73]:
```
# mean squared error for predictions
mse = mean_squared_error(y_test, y_pred)
```

[82]:
```
r2_score = ridge_regression.score(x_test, y_test)

f"mean squared error is {mse}, the model fits {r2_score * 100:.2f}% of the␣
↪inputs - {'OK' if r2_score > 0.6 else 'BAD'}"
```

[82]:
```
'mean squared error is 0.06318598222846789, the model fits 71.57% of the inputs
- OK'
```