Q1-1: Which of the following statement(s) is(are) TRUE about regularization parameter λ ?

A. *λ is the tuning parameter that decides how much we want to penalize the flexibility of our model.*
B. *λ is usually set using cross validation.*

1. True, True
2. True, False
3. False, True
4. False, False

Q1-1: Which of the following statement(s) is(are) TRUE about regularization parameter λ ?

A. *λ is the tuning parameter that decides how much we want to penalize the flexibility of our model.*

B. *λ is usually set using cross validation.*

1. True, True
2. True, False
3. False, True
4. False, False

- The optimization problem can be viewed as following:

$$\mathrm{minimize}(\mathrm{Loss}(\mathrm{Data}|\mathrm{Model}) + \lambda\,\mathrm{complexity}(\mathrm{Model}))$$

- If the regularization parameter is large then it requires a small model complexity
- We have learned how to use cross validate to set hyperparameters including regularization parameters.

Q1-2: Select the correct option about regression with L2 regularization (also called *Ridge Regression*).

A. *Ridge regression technique prevents coefficients from rising too high.*

B. *As $\lambda \to \infty$, the impact of the penalty grows, and the ridge regression coefficient estimates will approach infinity.*

1. Both statements are true.
2. Both statements are false.
3. Statement A is true, Statement B is false.
4. Statement B is true, Statement A is false.

Q1-2: Select the correct option about regression with L2 regularization (also called *Ridge Regression*).

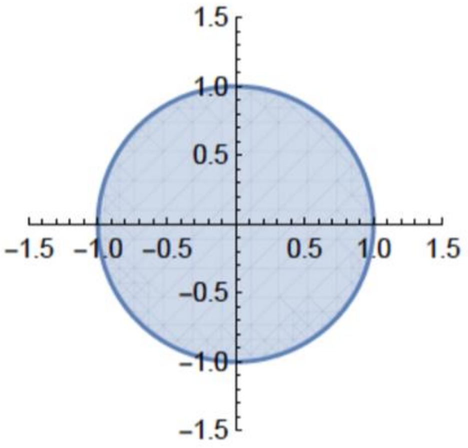
A. *Ridge regression technique prevents coefficients from rising too high.*

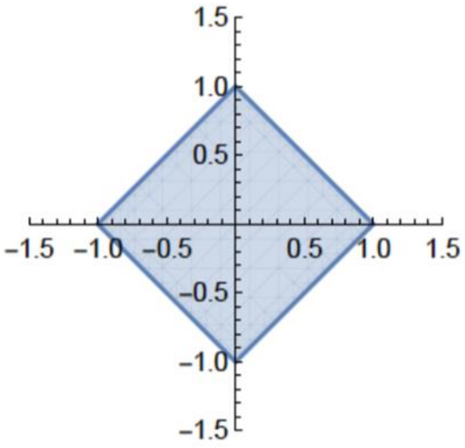
B. *As $\lambda \to \infty$, the impact of the penalty grows, and the ridge regression coefficient estimates will approach infinity.*

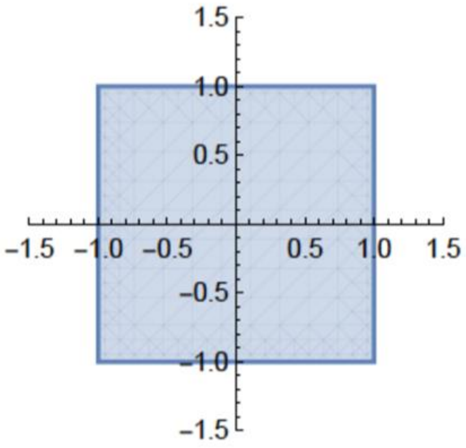
1. Both statements are true.
2. Both statements are false.
3. Statement A is true, Statement B is false.
4. Statement B is true, Statement A is false.

As $\lambda \to \infty$, the impact of the penalty grows, and the ridge regression coefficient estimates will approach zero.

Q1-3: Following figure shows 3-norm sketches: $||x||_p<1$ for p = 1, 2, ∞.
Recall that $||x||_\infty = \max\{|x_i| \text{ for all i}\}$

A: 

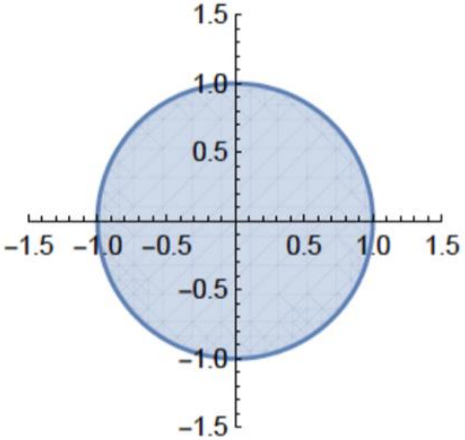
B: 

C: 

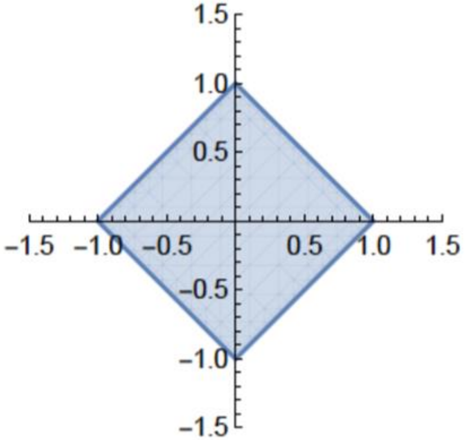
1. A: 1, B: 2, C: ∞
2. A: 2, B: 1, C: ∞
3. A: 2, B: ∞, C: 1
4. A: ∞, B: 2, C: 1

Q1-3: Following figure shows 3-norm sketches: $||x||_p < 1$ for p = 1, 2, ∞.
Recall that $||x||_\infty = \max\{|x_i| \text{ for all } i\}$

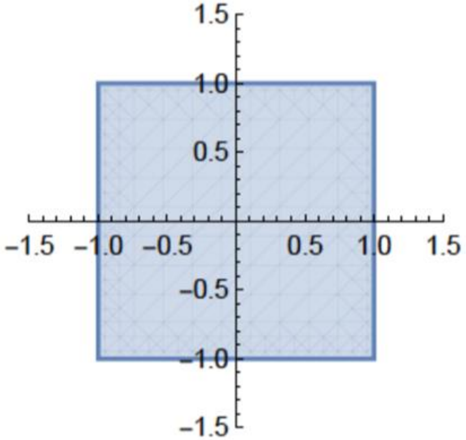
A:



B:



C:



1. A: 1, B: 2, C: ∞
2. A: 2, B: 1, C: ∞
3. A: 2, B: ∞, C: 1
4. A: ∞, B: 2, C: 1

Q2-1: Are these statements true or false?
(A) We need validation data to decide when to early stop.
(B) We can think early stopping as a regularization to limit the volume of parameter space reachable from the initial parameter.

1. True, True
2. True, False
3. False, True
4. False, False

Q2-1: Are these statements true or false?
(A) We need validation data to decide when to early stop.
(B) We can think early stopping as a regularization to limit the volume of parameter space reachable from the initial parameter.

1. True, True
2. True, False
3. False, True
4. False, False

(A) As is shown in the lecture.
(B) That's true. Early stopping will limit the training time and thus potentially limit the space the training can search.