# Activity 9

Damion Huppert

In [33]:
```
import numpy as np
import matplotlib.pyplot as plt
```

Question 1-a

$$X = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$$

$$y = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$$

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$Xw = \begin{bmatrix} w_1 + w_2 \\ -2w_1 - 2w_2 \end{bmatrix}$$

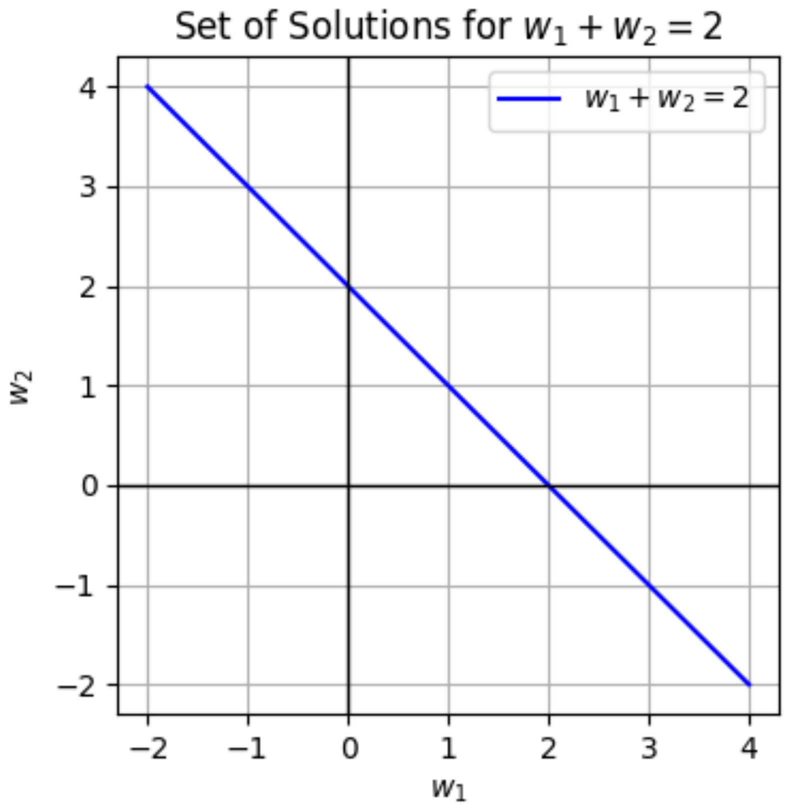
$$w_1 + w_2 = 2$$

$$w_1 + w_2 = 2$$

The solution is not unique because both equations are the same, the solution lies on the line $w_1 + w_2 = 2$

$$\min_w ||Xw - y||_2^2 = 0$$

In [34]:
```
w1 = np.linspace(-2, 4, 100)
w2 = 2 - w1

plt.figure(figsize=(4,4))
plt.plot(w1, w2, label=r'$w_1 + w_2 = 2$', color='blue')

plt.axhline(0, color='black', linewidth=1)
plt.axvline(0, color='black', linewidth=1)
plt.xlabel(r'$w_1$')
plt.ylabel(r'$w_2$')
plt.title('Set of Solutions for $w_1 + w_2 = 2$')
plt.legend()
plt.grid()
plt.show()
```




Question 1-b

$$\min_{w} ||w||_2^2 \quad \text{subject to} \quad w_1 + w_2 = 2$$

By analyzing the graph we can see $w$ will be smallest when $w = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

This solution is unique, because there is only one value that is the minimum of $||w||_2^2 = \sqrt{2}$

The error at the solution is 0 because (1, 1) lies on the line $w_1 + w_2 = 2$

Questoin 1-c

$$\hat{w} = \arg\min_{w} \left\{ ||Xw - y||_2^2 + \lambda ||w||_2^2 \right\}$$

$\|Xw - y\|_2^2 = (Xw - y)^T (Xw - y)$ and $\|w\|_2^2 = w^T w$

$$f(w) = (Xw - y)^T (Xw - y) + \lambda w^T w$$

$(Xw - y)^T = (Xw)^T - y^T$ and $y^T(Xw) = y(Xw)^T$

$$f(w) = (Xw)^T(Xw) - 2(Xw)^T y + y^T y + \lambda w^T w$$

$$f(w) = w^T X^T X w - 2 w^T X^T y + y^T y + \lambda w^T w$$

To find the minimum take the gradient and set it equal to 0

$$\nabla_w f(w) = 2X^T Xw - 2X^T y + 2\lambda w = 0$$

$$2X^T Xw + 2\lambda w = 2X^T y$$

$$X^T Xw + \lambda w = X^T y$$

$$(X^T X + \lambda I)w = X^T y$$

$$X^T X = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}$$

$$A = X^T X + \lambda I = \begin{bmatrix} 5+\lambda & 5 \\ 5 & 5+\lambda \end{bmatrix}$$

$$w = A^{-1} X^T y$$

$$A^{-1} = \begin{bmatrix} 5+\lambda & 5 \\ 5 & 5+\lambda \end{bmatrix}^{-1} = \frac{1}{25 + 10\lambda + \lambda^2 - 25} \begin{bmatrix} 5+\lambda & -5 \\ -5 & 5+\lambda \end{bmatrix}$$

$$A^{-1} = \frac{1}{\lambda(10+\lambda)} \begin{bmatrix} 5+\lambda & -5 \\ -5 & 5+\lambda \end{bmatrix}$$

$$w = \frac{1}{\lambda(10+\lambda)} \begin{bmatrix} 5+\lambda & -5 \\ -5 & 5+\lambda \end{bmatrix} \begin{bmatrix} 10 \\ 10 \end{bmatrix}$$

$$w = \frac{1}{\lambda(10+\lambda)} \begin{bmatrix} 10(5+\lambda) - 50 \\ 10(5+\lambda) - 50 \end{bmatrix}$$

$$w = \frac{10(5+\lambda) - 50}{\lambda(10+\lambda)} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$w = \frac{50 + 10\lambda - 50}{\lambda(10+\lambda)} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$w = \frac{10\lambda}{\lambda(10+\lambda)} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$w = \frac{10}{(10+\lambda)} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

In [35]:

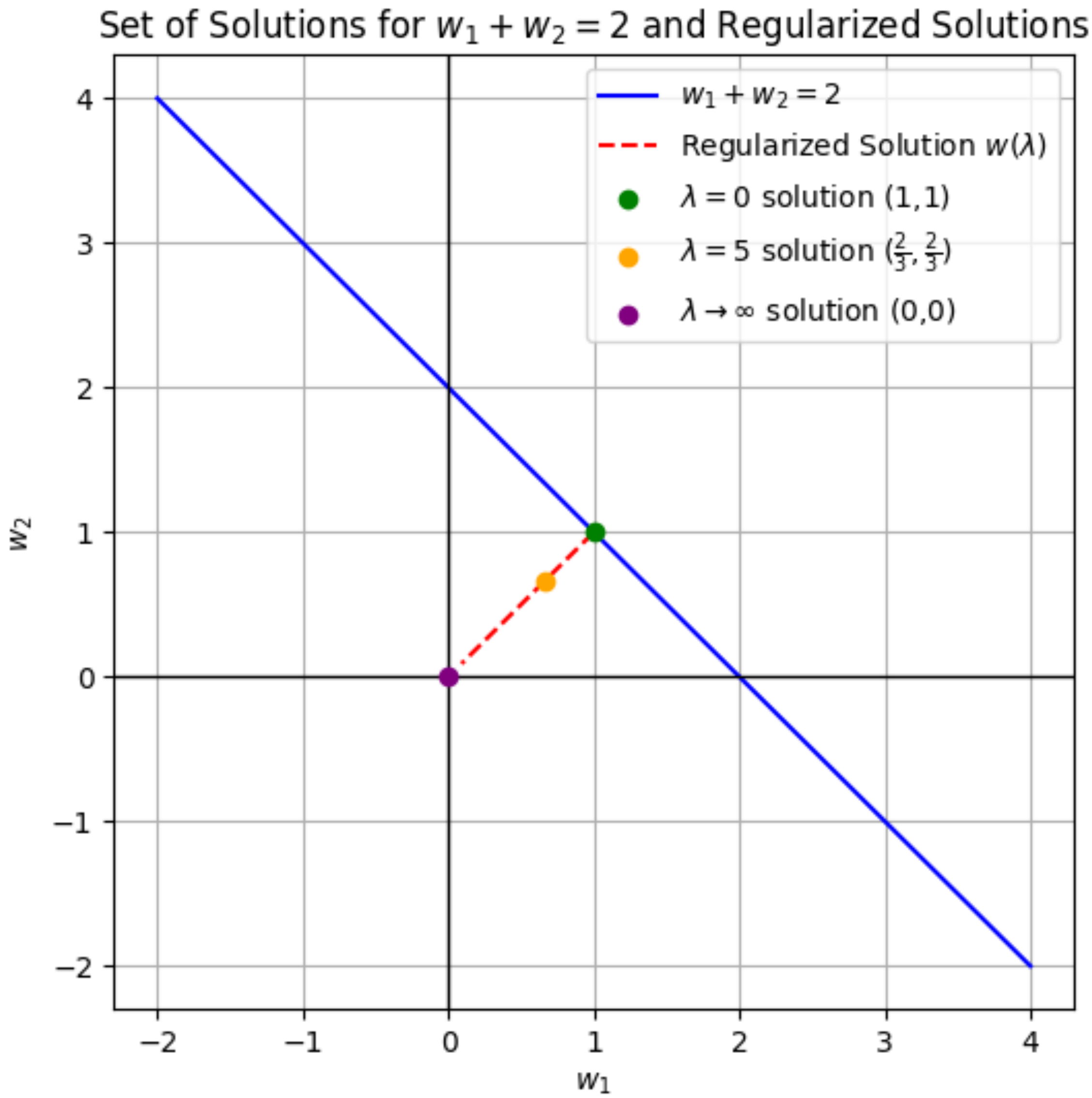
```
X = np.array([[1, 1], [-2, -2]])
y = np.array([2, -4])
print(X.T @ X)
print()
print(X.T @ y)
```

```
[[5 5]
 [5 5]]

[10 10]
```

### Questoin 1-d

In [36]:

```
w1_line = np.linspace(-2, 4, 100)
w2_line = 2 - w1_line

lambda_vals = np.linspace(0, 100, 100)

w1_reg = 10 / (10 + lambda_vals)
w2_reg = 10 / (10 + lambda_vals)

plt.figure(figsize=(6, 6))
plt.plot(w1_line, w2_line, label=r'$w_1 + w_2 = 2$', color='blue')

plt.plot(w1_reg, w2_reg, label=r'Regularized Solution $w(\lambda)$', color='

plt.scatter([1], [1], color='green', label=r'$\lambda = 0$ solution (1,1)',
plt.scatter([2/3], [2/3], color='orange', label=r'$\lambda = 5$ solution $(\
plt.scatter([0], [0], color='purple', label=r'$\lambda \to \infty$ solution

# Formatting
plt.axhline(0, color='black', linewidth=1)
plt.axvline(0, color='black', linewidth=1)
plt.xlabel(r'$w_1$')
plt.ylabel(r'$w_2$')
plt.title('Set of Solutions for $w_1 + w_2 = 2$ and Regularized Solutions')
plt.legend()
plt.grid()
plt.show()
```




Question 2-a

Columns of X:

$$c_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad c_2 = \begin{bmatrix} \gamma \\ -\gamma \\ -\gamma \\ \gamma \end{bmatrix}$$

$$c_1 \cdot c_2 = (1)(\gamma) + (1)(-\gamma) + (1)(-\gamma) + (1)(\gamma) = \gamma - \gamma - \gamma + \gamma = 0$$

Therefore because the dot product is always 0 $c_1$ and $c_2$ are orthoganal $\forall \gamma$

Question 2-b

$$X = \begin{bmatrix} 1 & \gamma \\ 1 & -\gamma \\ 1 & -\gamma \\ 1 & \gamma \end{bmatrix}$$

$$U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{31} \\ u_{41} & u_{42} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} s_{11} & 0 \\ 0 & s_{22} \end{bmatrix}$$

$$U \times \Sigma = \begin{bmatrix} u_{11}s_{11} & u_{12}s_{22} \\ u_{21}s_{11} & u_{22}s_{22} \\ u_{31}s_{11} & u_{31}s_{22} \\ u_{41}s_{11} & u_{42}s_{22} \end{bmatrix}$$

$$U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \\ u_{41} & u_{42} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} s_{11} & 0 \\ 0 & s_{22} \end{bmatrix}$$

From $X = U\Sigma$:

- $u_{11}s_{11} = 1$
- $u_{12}s_{22} = \gamma$
- $u_{21}s_{11} = 1$
- $u_{22}s_{22} = -\gamma$
- $u_{31}s_{11} = 1$
- $u_{32}s_{22} = -\gamma$
- $u_{41}s_{11} = 1$
- $u_{42}s_{22} = \gamma$

$$U = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}$$

Question 2-c

$$\min_w ||Xw - y||_2^2 = \min_w ||(U\Sigma)w - y||_2^2$$

$$w = \left((U\Sigma)^T(U\Sigma)\right)^{-1}(U\Sigma)^T y$$

$(U\Sigma)^T(U\Sigma) = \Sigma^T U^T U \Sigma$

$$w = \left(\Sigma^T U^T U \Sigma\right)^{-1}\Sigma^T U^T y$$

In [37]:
```
U = np.array([[1, 1], [1, -1], [1, -1], [1, 1]])
print(U.T @ U)
```

```
[[4 0]
 [0 4]]
```

Question 2-d

$$y = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$U^T U = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

$$w = \left(\Sigma^T U^T U \Sigma\right)^{-1}\Sigma^T U^T y$$

$$\Sigma^T U^T U \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4\gamma^2 \end{bmatrix}$$

$$(\Sigma^T U^T U \Sigma)^{-1} = \begin{bmatrix} 4 & 0 \\ 0 & 4\gamma^2 \end{bmatrix}^{-1} = \frac{1}{16\gamma^2}\begin{bmatrix} 4\gamma^2 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4\gamma^2} \end{bmatrix}$$

$$\Sigma^T U^T y = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2\gamma \end{bmatrix}$$

$$w = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4\gamma^2} \end{bmatrix} \begin{bmatrix} 2 \\ 2\gamma \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2\gamma} \end{bmatrix}$$

As $\gamma \to 0$ $w \to \begin{bmatrix} \frac{1}{2} \\ \infty \end{bmatrix}$

Question 2-e

$$\gamma = 0.1$$

$$\text{condition number} = \frac{1}{0.1} = 10$$

$$||w||_2^2 = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2 \times 0.1}\right)^2} = 5.025$$

$\gamma = 10^{-8}$

$$\text{condition number} = \frac{1}{10^{-8}} = 10^8$$

$$||w||_2^2 = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2 \times 10^{-8}}\right)^2} = 5 \times 10^7$$

Question 2-f

$$w_0 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2\gamma} \end{bmatrix}$$

$$w_\epsilon = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4\gamma^2} \end{bmatrix} \begin{bmatrix} 2+\epsilon \\ 2\gamma + \epsilon\gamma \end{bmatrix} = \begin{bmatrix} \frac{1}{2} + \frac{1}{4}\epsilon \\ \frac{2+\epsilon}{4\gamma} \end{bmatrix}$$

In [40]:
```
gamma = 0.1
epsilon = 0.01
w_0 = np.array([1/2, 1/(2*gamma)])
w_epsilon = np.array([[.5 + .25*epsilon],[(2+epsilon)/(4*gamma)]])
```

```python
print("gamma = 0.1")
print("w_0 =", w_0)
print("w_epsilon =", w_epsilon)

gamma = 10**-8
w_0 = np.array([1/2, 1/(2*gamma)])
w_epsilon = np.array([[.5 + .25*epsilon],[(2+epsilon)/(4*gamma)]])
print("gamma = 10^-8")
print("w_0 =", w_0)
print("w_epsilon =", w_epsilon)
```

```
gamma = 0.1
w_0 = [0.5 5. ]
w_epsilon = [[0.5025]
 [5.025 ]]
gamma = 10^-8
w_0 = [5.e-01 5.e+07]
w_epsilon = [[5.025e-01]
 [5.025e+07]]
```

Question 2-g

$$U = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}$$

$$y = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$y_\epsilon = \begin{bmatrix} 1+\epsilon \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$w = \arg\min_w ||Xw - y||_2^2 + \lambda ||w||_2^2$$

From 1-c:

$$w = (X^T X + \lambda I)^{-1} X^T y$$

$X = U\Sigma$

$$w = ((U\Sigma)^T U\Sigma + \lambda I)^{-1} (U\Sigma)^T y$$

$$w = \left(\Sigma^T \Sigma + \lambda I\right)^{-1} \Sigma^T U^T y.$$

$$w_0 = \left(\Sigma^T \Sigma + \lambda I\right)^{-1} \Sigma^T U^T y.$$

$$w_\epsilon = \left(\Sigma^T \Sigma + \lambda I\right)^{-1} \Sigma^T U^T y_\epsilon.$$

$$w_0 = \begin{bmatrix} \frac{2}{1+\lambda} \\ \frac{2\gamma}{\gamma^2+\lambda} \end{bmatrix}, \quad w_\epsilon = \begin{bmatrix} \frac{2+\epsilon}{1+\lambda} \\ \frac{2\gamma+\gamma\epsilon}{\gamma^2+\lambda} \end{bmatrix}$$

In [42]:

```
lambda_ = 0.1
gamma = 0.5
epsilon = 0.01

w_0 = np.array([
    2 / (1 + lambda_),
    (2 * gamma) / (gamma**2 + lambda_)
])

w_epsilon = np.array([
    (2 + epsilon) / (1 + lambda_),
    (2 * gamma + gamma * epsilon) / (gamma**2 + lambda_)
])

print("gamma = 0.1")
print("w_0 =", w_0)
print("w_epsilon =", w_epsilon)

epsilon = 10**-8
```

```python
w_0 = np.array([
    2 / (1 + lambda_),
    (2 * gamma) / (gamma**2 + lambda_)
])

w_epsilon = np.array([
    (2 + epsilon) / (1 + lambda_),
    (2 * gamma + gamma * epsilon) / (gamma**2 + lambda_)
])
print("gamma = 10^-8")
print("w_0 =", w_0)
print("w_epsilon =", w_epsilon)
```

```
gamma = 0.1
w_0 = [1.81818182 2.85714286]
w_epsilon = [1.82727273 2.87142857]
gamma = 10^-8
w_0 = [1.81818182 2.85714286]
w_epsilon = [1.81818183 2.85714287]
```

The values of $w_0$ and $w_\epsilon$ are now the same