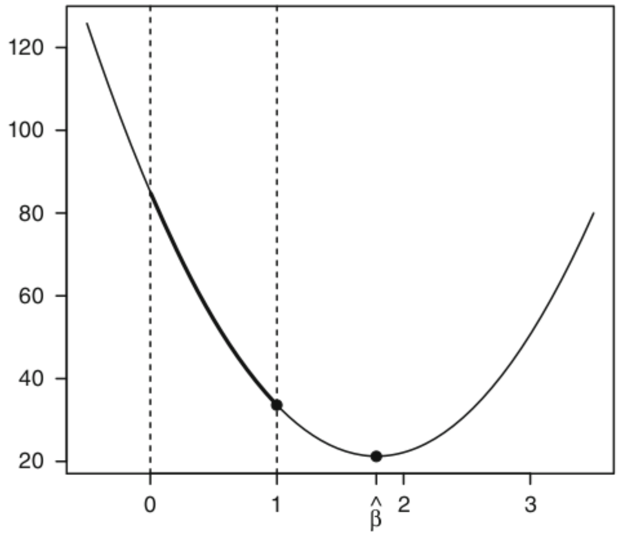
1.

$$\begin{aligned} &i)\ \mathbf{C} = (0, 5, -7, 0), && d = -2 \\ &ii)\ \mathbf{C} = (0, 2, -1, -1, 0), && d = 0 \end{aligned}$$

2. This plot shows the SSE in function of $\beta$:




The F-test score is $\dfrac{SSE_{H_0} - SSE}{SSE}$ where $SSE = SSE(\hat{\beta})$ that is the y in the plot where $x = \hat{\beta}$ and $SSE_{H_0} = SSE(\beta = 1)$
The plot for the numerator of Wald test is the same of the previous because the numerator is the same except for a square difference instead of normal difference.
The main difference between F-test and Wald test is the denominator in the first is $\hat{\theta}$ while in the second is $var(\hat{\theta})$. There is also a relation between test statistics $F$ and $W$: $W = rF$

3. The confidence interval for $\mu_0$ is $\mathbf{x}_0'\hat{\beta} \pm t_{n-p}(1-\alpha/2)\hat{\sigma}(\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}x_0)^{1/2}$ while for prediction is $\mathbf{x}_0'\hat{\beta} \pm t_{n-p}(1-\alpha/2)\hat{\sigma}(1+\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}x_0)^{1/2}$. The difference between the two formulas is the "1 +" near $x_0$ that become from the fact that the estimation try to give an confidence interval of a parameter , so with no variance, while the prediction try to estimate a aleatory variable with a expected value and a variance.

4. $f(x_0) = \mathbf{x}_0'\hat{\beta} \pm t_{n-2}(1-\alpha/2)\hat{\sigma}(1+\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}x_0)^{1/2}$ where $X = \begin{pmatrix} 1 & x_{11} \\ \vdots & \vdots \\ 1 & x_{n1} \end{pmatrix}$
and $p = 2$.
The length of interval is minimum when $\mathbf{x_0} = \mathbf{0}$

5. Increasing the number of parameters is not always a good idea: we can express the SPSE (expected squared prediction error as): $\underbrace{n\sigma^2}_{irreducible\ part} + \underbrace{|M|\sigma^2}_{Variance\ error} + \underbrace{\sum_{i=1}^{n}(u_{iM} - u_i)^2}_{BIAS^2}$

6. If we consider only $\mathbf{X_1}$ we have $\mathbf{X} = \mathbf{X_1}$ then

$$\begin{aligned} E\left[\tilde{\beta}_1\right] &= E\left[(\mathbf{X_1'X_1})^{-1}\mathbf{X_1'y}\right] \\ &= (\mathbf{X_1'X_1})^{-1}\mathbf{X_1'}E\left[\mathbf{y}\right] \\ &= (\mathbf{X_1'X_1})^{-1}\mathbf{X_1'}(\mathbf{X_1}\beta_{\mathbf{1}} + \mathbf{X_2}\beta_{\mathbf{2}}) \\ &= \beta_1 + (\mathbf{X_1'X_1})^{-1}\mathbf{X_1'X_2}\beta_{\mathbf{2}} \end{aligned}$$

The bias vanish if $\mathbf{X_1'X_2} = \mathbf{0}$ that in statical words mean that $\mathbf{X_1}$ and $\mathbf{X_2}$ are uncorrelated