MATH 367 - Week 3 Notes

Jasraj Sandhu

September 2023

Chain Rule

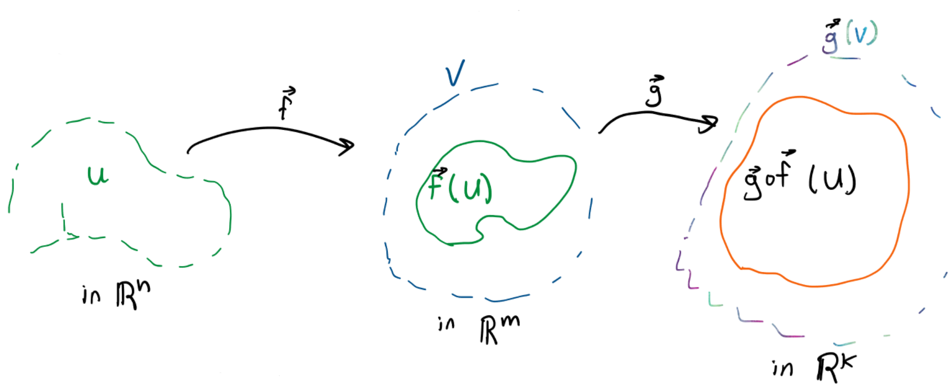
In the classical setting...

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0) \ .$$

In our general setting:

- If $\vec{f} : U \to \mathbb{R}^m$, where $U \subseteq \mathbb{R}^n$ is open, and $\vec{g} : V \to \mathbb{R}^k$, where $V \subseteq \mathbb{R}^m$ is open and $\vec{f}(U) \subseteq V$ (note that $\vec{f}(U)$ is the range of $\vec{f}$), then we may define $\vec{g} \circ \vec{f} : U \to \mathbb{R}^k$ by

$$\vec{g} \circ \vec{f}(\vec{x}_0) = \vec{g}(\vec{f}(\vec{x}_0)) \ .$$



When $\vec{f}$ is differentiable at $\vec{x}_0$ and $\vec{g}$ is differentiable at $\vec{f}(\vec{x}_0)$, then $\vec{g} \circ \vec{f}$ is differentiable at $\vec{x}_0$, and

$$D(\vec{g} \circ \vec{f})(\vec{x}_0) = D\vec{g}(\vec{f}(x_0)) \cdot D\vec{f}(\vec{x}_0) \ .$$

Note that $D\vec{g}(\vec{f}(x_0))$ is $k \times m$ and $D\vec{f}(\vec{x}_0)$ is $m \times n$, which means $D(\vec{g} \circ \vec{f})(\vec{x}_0)$ is $k \times n$.

1. Let $\vec{f}: \mathbb{R}^2 \to \mathbb{R}^2$ be defined as

$$\vec{f}(x,y) = \begin{bmatrix} xy \\ x+y \end{bmatrix}$$

and let $\vec{g}: \mathbb{R}^2 \to \mathbb{R}^3$ be defined as

$$\vec{g}(x,y) = \begin{bmatrix} xy^2 \\ 2y \\ x^2 \end{bmatrix} .$$

Use the chain rule to find $D(\vec{g} \circ \vec{f})(x,y)$.

Answer: Note that $\vec{g} \circ \vec{f}: \mathbb{R}^2 \to \mathbb{R}^3$. We know that

$$D(\vec{g} \circ \vec{f})(x,y) = D\vec{g}(\vec{f}(x,y)) \cdot D\vec{f}(x,y) .$$

So, first we evaluate $D\vec{f}(x,y)$. Since $\vec{f}$ is differentiable at (x,y) (it has to be based on the nature of the question), all partials for $\vec{f}$ exist, and so we can use the theorem from Week 2. Let $f_1 = xy$ and $f_2 = x+y$. Then

$$\begin{aligned} D\vec{f}(x,y) &= \begin{bmatrix} (f_1)_x & (f_1)_y \\ (f_2)_x & (f_2)_y \end{bmatrix} \\ &= \begin{bmatrix} (xy)_x & (xy)_y \\ (x+y)_x & (x+y)_y \end{bmatrix} \\ &= \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix} . \end{aligned}$$

Next, we evaluate $\vec{g}(x,y)$. Just like $\vec{f}$, since $\vec{g}$ is differentiable at (x,y), all partials for $\vec{g}$ exist, and so we can use the theorem from Week 2. Let $g_1 = xy^2$, $g_2 = 2y$, and $g_3 = x^2$. Then

$$\begin{aligned} D\vec{g}(x,y) &= \begin{bmatrix} (g_1)_x & (g_1)_y \\ (g_2)_x & (g_2)_y \\ (g_3)_x & (g_3)_y \end{bmatrix} \\ &= \begin{bmatrix} (xy^2)_x & (xy^2)_y \\ (2y)_x & (2y)_y \\ (x^2)_x & (x^2)_y \end{bmatrix} \\ &= \begin{bmatrix} y^2 & 2xy \\ 0 & 2 \\ 2x & 0 \end{bmatrix} . \end{aligned}$$

Then applying the chain rule gives us

$$\begin{aligned} D(\vec{g} \circ \vec{f})(x,y) &= D\vec{g}(\vec{f}(x,y)) \cdot D\vec{f}(x,y) \\ &= D\vec{g}(xy, x+y) \cdot \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} y^2 & 2xy \\ 0 & 2 \\ 2x & 0 \end{bmatrix}\Bigg|_{(xy,x+y)} \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} (x+y)^2 & 2(xy)(x+y) \\ 0 & 2 \\ 2(xy) & 0 \end{bmatrix} \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} (x+y)^2 & 2xy(x+y) \\ 0 & 2 \\ 2xy & 0 \end{bmatrix} \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} y(x+y)^2 + 2xy(x+y) & x(x+y)^2 + 2xy(x+y) \\ 2 & 2 \\ 2xy^2 & 2x^2y \end{bmatrix} .
\end{aligned}
$$

2. Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ is differentiable with $x = x(t)$, $y = y(t)$, and $z = z(t)$ (all differentiable w.r.t. to t). Define $F(t) = f(x(t), y(t), z(t))$. Find an expression for $\dfrac{dF}{dt}$.

Let $\vec{r}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}$ so that $F(t) = f(\vec{r}(t))$. Then

$$DF(t) = Df(\vec{r}(t))D\vec{r}(t) \ ,$$

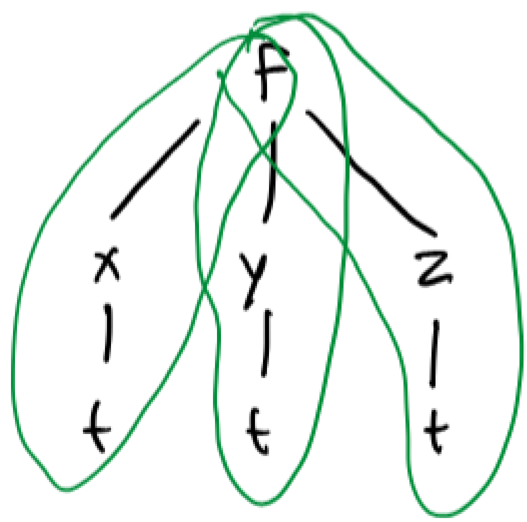
where $DF(t)$ is 1×1, since $Df(\vec{r}(t))$ is 1×3 and $D\vec{r}(t)$ is 3×1. This implies that

$$\frac{dF}{dt} = \begin{bmatrix} f_x(\vec{r}(t)) & f_y(\vec{r}(t)) & f_z(\vec{r}(t)) \end{bmatrix} \begin{bmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \\ \dfrac{dz}{dt} \end{bmatrix} \ .$$

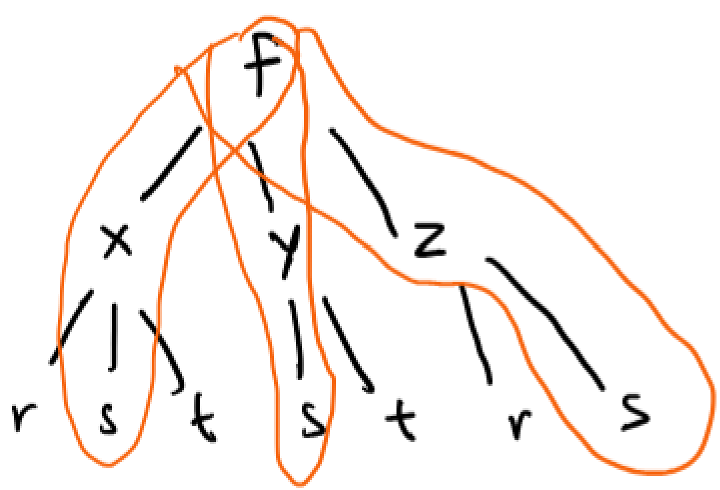
Thus, an expression for $\dfrac{dF}{dt}$ is

$$\boxed{\begin{aligned} \frac{dF}{dt} = \ & f_x(x(t), y(t), z(t))\frac{dx}{dt} + f_y(x(t), y(t), z(t))\frac{dy}{dt} \\ & + f_z(x(t), y(t), z(t))\frac{dz}{dt} \end{aligned}}$$

- A **dependency tree** can be used to find such formulae:



- To find $\dfrac{dF}{dt}$, add terms for each branch terminating in t.



Then in this case,

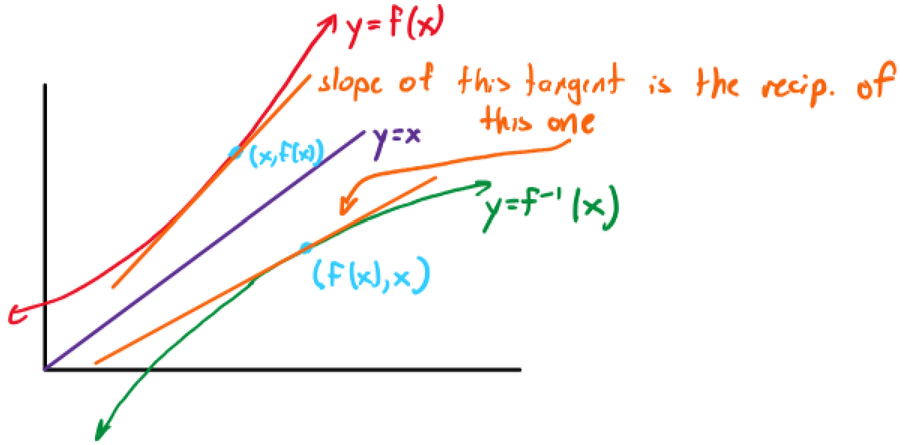
$$\frac{\partial F}{\partial s} = f_x \cdot \frac{\partial x}{\partial s} + f_y \cdot \frac{\partial y}{\partial s} + f_z \cdot \frac{\partial z}{\partial s} \ .$$

Inverse Function Theorem

Classical Inverse Function Theorem

If f is **invertible** on an interval I (i.e. there exists a function g such that $(f \circ g)(x) = x = (g \circ f)(x)$) and differentiable with $f'(x) \neq 0$ on I, then f^{-1} is differentiable on $f(I)$ and

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)} \ .$$



Another motivation:

- Let A be $n \times n$ and invertible such that

  $$\vec{f}(\vec{x}) = A\vec{x} + \vec{b}$$

  is invertible. Then

  $$\vec{f}^{-1}(\vec{x}) = A^{-1}(\vec{x} - \vec{b}) \ .$$

  (One can check that $(\vec{f}^{-1} \circ \vec{f})(\vec{x}) = \vec{x}$, etc).

- We know that $\vec{f}$ and $\vec{f}^{-1}$ are affine and therefore differentiable with $D\vec{f} = A$ and $D\vec{f}^{-1} = A^{-1}$.

Theorem (Inverse Function Theorem)

Suppose $\vec{f} : U \to \mathbb{R}^n$, where $U \subseteq \mathbb{R}^n$, is differentiable at $\vec{x}_0$ and that $D\vec{f}(\vec{x}_0)$, which is $m \times n$ (mistake in notes? I think it should say $D\vec{F}(\vec{x}_0)$ is $n \times n$, since you can't find the inverse of a non-square matrix), is an invertible matrix. Then $\vec{f}$ has a **local** inverse around the point $\vec{x}_0$, call it $\vec{g}$, and

$$D\vec{g}(\vec{f}(\vec{x}_0)) = D\vec{f}(\vec{x}_0)^{-1}.$$

Note that **local** here refers to this: $\exists$ an open set V such that $\vec{x}_0 \in V$, with $\vec{g} \circ \vec{f}(\vec{x}) = \vec{x}$ for all $\vec{x} \in V$.

1. Let $\vec{f}(x, y) = \begin{bmatrix} 2x - y \\ x^2 - y^2 \end{bmatrix}$.

   (a) For which pairs (x_0, y_0) do the hypotheses hold for in the Inverse Function Theorem?

   Let $f_1 = 2x - y$ and $f_2 = x^2 - y^2$. Then

$$\begin{aligned} D\vec{f}(x, y) &= \begin{bmatrix} (f_1)_x & (f_1)_y \\ (f_2)_x & (f_2)_y \end{bmatrix} \\ &= \begin{bmatrix} (2x - y)_x & (2x - y)_y \\ (x^2 - y^2)_x & (x^2 - y^2)_y \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 \\ 2x & -2y \end{bmatrix} . \end{aligned}$$

   We have what we need provided this matrix is invertible. Now,

$$\begin{aligned} \det\left(D\vec{f}\right) &= \det \begin{bmatrix} 2 & -1 \\ 2x & -2y \end{bmatrix} \\ &= (2)(-2y) - (-1)(2x) \\ &= -4y + 2x . \end{aligned}$$

   So, for any pair (x_0, y_0) with $-4y_0 + 2x_0 \neq 0$, the Inverse Function Theorem applies (a matrix is invertible if $\det(A) \neq 0$) Note that this is everything off of the line $y = \dfrac{x}{2}$.

(b) Find $D\vec{g}(\vec{f}(3,4))$, where $\vec{g}$ is the local inverse to $\vec{f}$ at $(3,4)$ given in the theorem.

We know that

$$\begin{aligned} D\vec{f}(3,4) &= \begin{bmatrix} 2 & -1 \\ 2(3) & -2(4) \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 \\ 6 & -8 \end{bmatrix} . \end{aligned}$$

Then by the Inverse Function Theorem, since $\vec{g}$ is the local inverse to f at $(3,4)$, it follows that

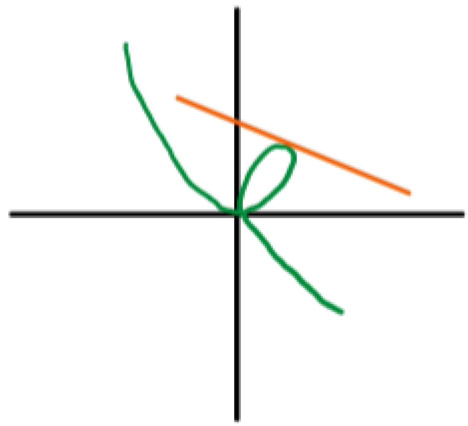
$$\begin{aligned} D\vec{g}(\vec{f}(3,4)) &= D\vec{f}(3,4)^{-1} \\ &= \begin{bmatrix} 2 & -1 \\ 6 & -8 \end{bmatrix}^{-1} \\ &= \frac{1}{(2)(-8)-(-1)(6)} \begin{bmatrix} -8 & 1 \\ -6 & 2 \end{bmatrix} \\ &= \frac{1}{-16+6} \begin{bmatrix} -8 & 1 \\ -6 & 2 \end{bmatrix} \\ &= -\frac{1}{10} \begin{bmatrix} -8 & 1 \\ -6 & 2 \end{bmatrix} . \end{aligned}$$

- If we write $\vec{f}(x,y) = \begin{bmatrix} u(x,y) \\ v(x,y) \end{bmatrix}$, then $\vec{g}(u,v) = \begin{bmatrix} x(u,v) \\ y(u,v) \end{bmatrix}$ is a function of u, v.

- $\dfrac{\partial x}{\partial u}$ is the $1,1$ entry of $Dg(u,v)$, etc...

Implicit Functions

- Recall for a curve like $x^3 + y^3 = 2xy$ (Folium),



  we may still compute tangent lines at specific points even though it's not a function. This is because the curve is **locally a function** (zoom in far enough to pass the vertical line test).

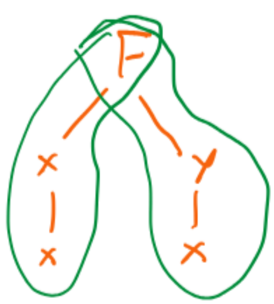
- We may assume $y = y(x)$ and then consider

$$F(x) = x^3 + y(x)^3 - 2x \cdot y(x) = 0 \ .$$

  Note that $F'(x) = \dfrac{dF}{dx} = \dfrac{d}{dx}[0] = 0$. Then by the chain rule,

$$\begin{aligned} 0 &= \frac{\partial F}{\partial x} \\ &= (x^3 + y^3 - 2xy)_x x_x + (x^3 + y^3 - 2xy)_y y_x \\ &= (x^3 + y^3 - 2xy)_x \cdot 1 + (x^3 + y^3 - 2xy)_y \frac{dy}{dx} \ . \end{aligned}$$

  Then from the dependency tree,



we get that

$$
\begin{aligned}
0 &= (3x^2 - 2y) + (3y^2 - 2x)\frac{dy}{dx} \\
\frac{dy}{dx} &= \frac{2y - 3x^2}{3y^2 - 2x} \ .
\end{aligned}
$$

- In general, for any differential function $F : U \to \mathbb{R}$ with

  $$F(x_1, \ldots, x_n) = 0 \ ,$$

  we have that

  $$\frac{\partial x_i}{\partial x_j} = -\frac{F_{x_j}}{F_{x_i}}$$

Theorem (Implicit Function Theorem)

Suppose $f_1, \ldots . f_n$ are differentiable scalar functions in the variables $y_1, \ldots, y_n, x_1, \ldots, x_m$, and consider the equations

$$\begin{aligned} f_1(b_1, \ldots, b_n, a, \ldots, a_m) &= c_1 \\ &\vdots \\ f_1(b_1, \ldots, b_n, a, \ldots, a_m) &= c_n \ , \end{aligned}$$

where $c_1, \ldots, c_n$ are constants. Write

$$\vec{F} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

so that $\vec{F}$ is a function from $U \subseteq \mathbb{R}^{m+n}$ into $\mathbb{R}^n$ (i.e. $\vec{F} : U \to \mathbb{R}^n$), and

$$D\vec{F}(b_1, \ldots, b_n, a_1, \ldots, a_m) = \begin{bmatrix} B \mid A \end{bmatrix} \ .$$

If B is invertible, then we can locally solve for $y_1, \ldots, y_n$ in terms of $x_1, \ldots, x_m$ as

$$\vec{g}(x_1, \ldots, x_m) = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

and

$$D\vec{g}(a_1, \ldots, a_m) = -B^{-1}A \ ,$$

where $-B^{-1}A$ is $n \times m$.

1. Suppose we were given

$$x^2 - y + z - uv = -2$$
$$2x - y^3 + zu^2 = -2$$

on $(x, y, z, u, v) = (1, 0, -1, 2, 1)$,

(a) Determine if x, y can be solved in terms of z, u, v.

To use the implicit function theorem, define

$$\vec{F}(x, y, z, u, v) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} x^2 - y + z - uv \\ 2x - y^3 + zu^2 \end{bmatrix}$$

such that $\vec{F} : U \to \mathbb{R}^2$, where $U \subseteq \mathbb{R}^{m+n}$. Here $n = 2$ since the codomain is $\mathbb{R}^2$ and the domain U is a subset of $\mathbb{R}^5$, which means that

$$\begin{aligned} m + n &= 5 \\ m + 2 &= 5 \\ m &= 3 \; . \end{aligned}$$

Now, the derivative of F is given by

$$\begin{aligned} D\vec{F}(x, y, z, u, v) &= \begin{bmatrix} (f_1)_x & (f_1)_y & (f_1)_z & (f_1)_u & (f_1)_v \\ (f_2)_x & (f_2)_y & (f_2)_z & (f_2)_u & (f_2)_v \end{bmatrix} \\ &= \begin{bmatrix} 2x & -1 & 1 & -v & -u \\ 2 & -3y^2 & u^2 & 2zu & 0 \end{bmatrix} \; . \end{aligned}$$

Then by the implicit function theorem,

$$\begin{aligned} D\vec{F}(1, 0, -1, 2, 1) &= \begin{bmatrix} 2(1) & -1 & 1 & -1 & -2 \\ 2 & -3(0)^2 & 2^2 & 2(-1)(2) & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 & 1 & -1 & -2 \\ 2 & 0 & 4 & -4 & 0 \end{bmatrix} \\ &= \begin{bmatrix} B \mid A \end{bmatrix} \; , \end{aligned}$$

where B is the $n \times n$ matrix consisting of the entries of the x and y column (column 1 and column 2, respectively) given by

$$B = \begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}$$

and A is the $n \times m$ matrix

$$A = \begin{bmatrix} 1 & -1 & -2 \\ 4 & -4 & 0 \end{bmatrix} \; .$$

Note that B is invertible since

$$\det(B) = (2 \cdot 0) - (-1 \cdot 2) = 0 - (-2) = 2 \neq 0 \; .$$

Then since B is invertible, we can define the implicit function

$$\vec{g}(z, u, v) = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x(z, u, v) \\ y(z, u, v) \end{bmatrix} ,$$

where x, y are the variables we want to solve in terms of z, u, v. Notice that the variables we want to solve **in terms of** are inputs to this new function $\vec{g}$, while the variables we want to solve for are the components of the output of this vector-valued function $\vec{g}$, written as functions of the variables we are solving in terms of.

(b) Can z, u be solved in terms of x, y, v?

From part (a), we found that

$$D\vec{F}(1, 0, -1, 2, 1) = \begin{bmatrix} 2 & -1 & 1 & -1 & -2 \\ 2 & 0 & 4 & -4 & 0 \end{bmatrix} .$$

Note that column 3 and column 4 correspond the the variables z and u, respectively. Then in this case, B is the $n \times n$ matrix

$$B = \begin{bmatrix} 1 & -1 \\ 4 & -4 \end{bmatrix} .$$

Since $\det(B) = 0$, this means that we cannot use the implicit function theorem. Thus, we cannot solve for z, u in terms of x, y, v.

2. Show that x, y, z may be solved in terms of u in

$$\begin{aligned} x^2 + z &= 2 \\ yz^2 &= -1 \\ x + zu - e^u &= 0 \end{aligned}$$

at $(x, y, z, u) = (1, -1, 1, 0)$.

Define

$$\vec{F}(x, y, z, u) = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} x^2 + z \\ yz^2 \\ x + zu - e^u \end{bmatrix} .$$

such that $\vec{F} : U \to \mathbb{R}^n$, where $U \subseteq \mathbb{R}^{m+n}$. Here $n = 3$ since the codomain is $\mathbb{R}^3$, and $m + n = 4$ since the domain is $\mathbb{R}^4$, which means that

$$\begin{aligned} m + n &= 4 \\ m + 3 &= 4 \\ m &= 1 . \end{aligned}$$

Now, the derivative of $\vec{F}(x, y, z, u)$ is given by

$$\begin{aligned} D\vec{F}(x, y, z, u) &= \begin{bmatrix} (f_1)_x & (f_1)_y & (f_1)_z & (f_1)_u \\ (f_2)_x & (f_2)_y & (f_2)_z & (f_2)_u \\ (f_3)_x & (f_3)_y & (f_3)_z & (f_3)_u \end{bmatrix} \\ &= \begin{bmatrix} 2x & 0 & 1 & 0 \\ 0 & z^2 & 2yz & 0 \\ 1 & 0 & u & z - e^u \end{bmatrix} . \end{aligned}$$

Then

$$\begin{aligned} D\vec{F}(1, -1, 1, 0) &= \begin{bmatrix} 2(1) & 0 & 1 & 0 \\ 0 & (1)^2 & 2(-1)(1) & 0 \\ 1 & 0 & 0 & 1 - e^0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

where B is the $n \times n$ matrix consisting of the columns corresponding to the variables x, y, z (the variables we are trying to solve for), which is given by

$$B = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & -2 \\ 1 & 0 & 0 \end{bmatrix} .$$

Thus, since

$$\begin{aligned}
\det(B) &= \det\left(\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & -2 \\ 1 & 0 & 0 \end{bmatrix}\right) \\
&= 1[(0 \cdot (-2)) - (1 \cdot 1)] - 0 + 0 \\
&= 1(0 - 1) \\
&= 1(-1) \\
&= -1 \\
&\neq 0 \ ,
\end{aligned}$$

B is invertible which means that the inverse theorem applies.