MA 229/MA 235 - Lecture 3

IISc

Recap

Recap

- First and second derivative tests.

Recap

- First and second derivative tests.
- Proof of Clairaut's theorem.

Recap

- First and second derivative tests.
- Proof of Clairaut's theorem.
- Proof of Taylor's theorem (integral form of remainder).

Review of multivariable calculus: Positive-definiteness

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f.

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \; \forall \; v \neq 0$.

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \ \forall \ v \neq 0$. Often, optimization studies convex functions.

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \ \forall \ v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e.,

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \ \forall \ v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$,

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \; \forall \; v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$, it is said to be *positive-definite* if $\sum_{i,j} A_{ij} v_i v_j > 0$ for all $v \neq 0$, i.e.,

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \; \forall \; v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$, it is said to be *positive-definite* if $\sum_{i,j} A_{ij} v_i v_j > 0$ for all $v \neq 0$, i.e., $v^T A v > 0$ for all $v \neq 0$.

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \; \forall \; v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$, it is said to be *positive-definite* if $\sum_{i,j} A_{ij} v_i v_j > 0$ for all $v \neq 0$, i.e., $v^T A v > 0$ for all $v \neq 0$.
- This condition is somewhat subtle.

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \ \forall \ v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$, it is said to be *positive-definite* if $\sum_{i,j} A_{ij} v_i v_j > 0$ for all $v \neq 0$, i.e., $v^T A v > 0$ for all $v \neq 0$.
- This condition is somewhat subtle. For 2×2 matrices, $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, supposing $v^T = (x, y)$, we see that

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \; \forall \; v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$, it is said to be *positive-definite* if $\sum_{i,j} A_{ij} v_i v_j > 0$ for all $v \neq 0$, i.e., $v^T A v > 0$ for all $v \neq 0$.
- This condition is somewhat subtle. For 2×2 matrices, $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, supposing $v^T = (x, y)$, we see that $v^T A v = ax^2 + dy^2 + 2bxy$.

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \ \forall \ v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$, it is said to be *positive-definite* if $\sum_{i,j} A_{ij} v_i v_j > 0$ for all $v \neq 0$, i.e., $v^T A v > 0$ for all $v \neq 0$.
- This condition is somewhat subtle. For 2×2 matrices, $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, supposing $v^T = (x, y)$, we see that $v^T A v = ax^2 + dy^2 + 2bxy$. This expression is positive for all $(x, y) \neq (0, 0)$ if and only $a > 0$ and $ad - b^2 = \det(A) > 0$. (Why ?)

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \ \forall \ v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$, it is said to be *positive-definite* if $\sum_{i,j} A_{ij} v_i v_j > 0$ for all $v \neq 0$, i.e., $v^T A v > 0$ for all $v \neq 0$.
- This condition is somewhat subtle. For 2×2 matrices, $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, supposing $v^T = (x, y)$, we see that $v^T A v = ax^2 + dy^2 + 2bxy$. This expression is positive for all $(x, y) \neq (0, 0)$ if and only $a > 0$ and $ad - b^2 = \det(A) > 0$. (Why ?)
- Similar but more complicated conditions exist for $n \times n$ matrices. (

Review of multivariable calculus: Positive-definiteness

- The matrix f_{ij} is called the 'Hessian' matrix of f. For future purposes a twice differentiable function whose second partials are continuous is said to be convex if $\sum_{i,j} v_i f_{ij}(a) v_j \geq 0 \; \forall \; v \neq 0$. Often, optimization studies convex functions.
- In general, given a symmetric $n \times n$ real matrix A, i.e., $A^T = A$, i.e., $A_{ij} = A_{ji}$, it is said to be *positive-definite* if $\sum_{i,j} A_{ij} v_i v_j > 0$ for all $v \neq 0$, i.e., $v^T A v > 0$ for all $v \neq 0$.
- This condition is somewhat subtle. For 2×2 matrices, $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, supposing $v^T = (x, y)$, we see that $v^T A v = ax^2 + dy^2 + 2bxy$. This expression is positive for all $(x, y) \neq (0, 0)$ if and only $a > 0$ and $ad - b^2 = \det(A) > 0$. (Why ?)
- Similar but more complicated conditions exist for $n \times n$ matrices. (Positive-definiteness is the same as having only positive eigenvalues by the way.)

Review of multivariable calculus: Example

- Find local extrema of $f(x,y,z) = x^2 - y^2$ on $\mathbb{R}^2$:
  $\nabla f = (2x, -2y) = (0,0)$ only at the origin.

- Find local extrema of $f(x, y, z) = x^2 - y^2$ on $\mathbb{R}^2$:
  $\nabla f = (2x, -2y) = (0, 0)$ only at the origin. The Hessian is
  $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$.

Review of multivariable calculus: Example

- Find local extrema of $f(x, y, z) = x^2 - y^2$ on $\mathbb{R}^2$:
  $\nabla f = (2x, -2y) = (0, 0)$ only at the origin. The Hessian is
  $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$.
- So the origin is neither a local max nor a local min (despite the second derivative being non-zero ! (

Review of multivariable calculus: Example

- Find local extrema of $f(x, y, z) = x^2 - y^2$ on $\mathbb{R}^2$:
  $\nabla f = (2x, -2y) = (0, 0)$ only at the origin. The Hessian is
  $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$.
- So the origin is neither a local max nor a local min (despite the second derivative being non-zero ! (and none of the eigenvalues being zero)).

Review of multivariable calculus: Example

- Find local extrema of $f(x, y, z) = x^2 - y^2$ on $\mathbb{R}^2$:
$\nabla f = (2x, -2y) = (0, 0)$ only at the origin. The Hessian is
$\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$.
- So the origin is neither a local max nor a local min (despite the second derivative being non-zero ! (and none of the eigenvalues being zero)).
- There is a direction in which f increases and a direction in which it decreases.

Review of multivariable calculus: Example

- Find local extrema of $f(x, y, z) = x^2 - y^2$ on $\mathbb{R}^2$:
  $\nabla f = (2x, -2y) = (0, 0)$ only at the origin. The Hessian is
  $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$.
- So the origin is neither a local max nor a local min (despite the second derivative being non-zero ! (and none of the eigenvalues being zero)).
- There is a direction in which f increases and a direction in which it decreases. Such points are called "Saddle points".

Review of multivariable calculus: Example

- Find local extrema of $f(x, y, z) = x^2 - y^2$ on $\mathbb{R}^2$:
  $\nabla f = (2x, -2y) = (0, 0)$ only at the origin. The Hessian is
  $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$.
- So the origin is neither a local max nor a local min (despite the second derivative being non-zero ! (and none of the eigenvalues being zero)).
- There is a direction in which f increases and a direction in which it decreases. Such points are called "Saddle points".
- More generally, the eigenvectors and eigenvalues of the Hessian tell us about these "principal" directions.

The Einstein summation convention

The Einstein summation convention

- It is painful

The Einstein summation convention

- It is painful to keep using Σ

The Einstein summation convention

- It is painful to keep using Σ every time we want to

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over.

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- “Sum over indices that are repeated

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- “Sum over indices that are repeated above and below.

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector,

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector,

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w\rangle = v_i w^i$.

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix,

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector,

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w\rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space,

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w\rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted with *subscripts* (

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w\rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted with *subscripts* (like $e_1, e_2, \ldots$) and

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted with *subscripts* (like $e_1, e_2, \ldots$) and components w.r.t a basis are

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w\rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted with *subscripts* (like $e_1, e_2, \dots$) and components w.r.t a basis are denoted with superscripts (

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted with *subscripts* (like $e_1, e_2, \ldots$) and components w.r.t a basis are denoted with superscripts (like $v = v^i e_i$).

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted with *subscripts* (like $e_1, e_2, \ldots$) and components w.r.t a basis are denoted with superscripts (like $v = v^i e_i$). For the dual space,

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w\rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted with *subscripts* (like $e_1, e_2, \ldots$) and components w.r.t a basis are denoted with superscripts (like $v = v^i e_i$). For the dual space, the indices are flipped, i.e.,

The Einstein summation convention

- It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us.
- "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
- Examples:
  1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w \rangle = v_i w^i$.
  2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
  3. $Tr(A) = A^i_i$.
  4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
  5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.
- If V is a vector space, basis vectors are denoted with *subscripts* (like $e_1, e_2, \ldots$) and components w.r.t a basis are denoted with superscripts (like $v = v^i e_i$). For the dual space, the indices are flipped, i.e., $e_1^* = e^1, \ldots$ and $\omega = \omega_i e^i$.

Solving equations implicitly

Solving equations implicitly

- Consider $x^2 + y^2 = 1$.

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x?

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$.

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x!

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle,

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$.

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$,

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).
- What about $x^{\sin(y)} + y^2 + e^{xy} = 1$?

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).
- What about $x^{\sin(y)} + y^2 + e^{xy} = 1$? It is not clear

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).
- What about $x^{\sin(y)} + y^2 + e^{xy} = 1$? It is not clear whether we can solve

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).
- What about $x^{\sin(y)} + y^2 + e^{xy} = 1$? It is not clear whether we can solve for either variable

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).
- What about $x^{\sin(y)} + y^2 + e^{xy} = 1$? It is not clear whether we can solve for either variable in terms of the other.

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).
- What about $x^{\sin(y)} + y^2 + e^{xy} = 1$? It is not clear whether we can solve for either variable in terms of the other. Even if we can,

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).
- What about $x^{\sin(y)} + y^2 + e^{xy} = 1$? It is not clear whether we can solve for either variable in terms of the other. Even if we can, it is unlikely that

Solving equations implicitly

- Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not!
- $y = \pm\sqrt{1 - x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$.
- The best we can say is “near” (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y.
- We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x).
- What about $x^{\sin(y)} + y^2 + e^{xy} = 1$? It is not clear whether we can solve for either variable in terms of the other. Even if we can, it is unlikely that we can write a simple formula for the solution.

Linear equations

Linear equations

- More generally,

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$,

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask,

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions,

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where A, B are constant matrices (of what type?).

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where A, B are constant matrices (of what type?).
- So if $Ax + By = c$,

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where A, B are constant matrices (of what type?).
- So if $Ax + By = c$, can we solve for y in terms of x, c uniquely?

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where A, B are constant matrices (of what type?).
- So if $Ax + By = c$, can we solve for y in terms of x, c uniquely? $By = c - Ax$.

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where A, B are constant matrices (of what type?).
- So if $Ax + By = c$, can we solve for y in terms of x, c uniquely? $By = c - Ax$. Therefore, this question

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where A, B are constant matrices (of what type?).
- So if $Ax + By = c$, can we solve for y in terms of x, c uniquely? $By = c - Ax$. Therefore, this question has an affirmative answer

Linear equations

- More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?"
- Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where A, B are constant matrices (of what type?).
- So if $Ax + By = c$, can we solve for y in terms of x, c uniquely? $By = c - Ax$. Therefore, this question has an affirmative answer iff B is an invertible matrix.

A special case - the inverse function theorem

A special case - the inverse function theorem

- Often in maths,

A special case - the inverse function theorem

- Often in maths, general cases can be

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e.,

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$,

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c?

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is,

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e.,

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism?

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$,

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$, by monotonicity,

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$, by monotonicity, this problem is not hard.

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$, by monotonicity, this problem is not hard.
- In general,

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$, by monotonicity, this problem is not hard.
- In general, we need a theorem.

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$, by monotonicity, this problem is not hard.
- In general, we need a theorem. We *expect* that

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$, by monotonicity, this problem is not hard.
- In general, we need a theorem. We *expect* that it is good enough for

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$, by monotonicity, this problem is not hard.
- In general, we need a theorem. We *expect* that it is good enough for DF_a to be invertible.

A special case - the inverse function theorem

- Often in maths, general cases can be reduced to the simplest non-trivial special cases.
- Suppose $m = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}?
- When $n = 1$, by monotonicity, this problem is not hard.
- In general, we need a theorem. We *expect* that it is good enough for DF_a to be invertible. Indeed, this expectation is true.

A special case - the inverse function theorem (IFT)

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets.

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞).

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$,

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism.

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover,

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$.

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case,

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.
- The proof of this theorem

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.
- The proof of this theorem involves the contraction mapping principle. (

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.
- The proof of this theorem involves the contraction mapping principle. (Optional fun fact:

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.
- The proof of this theorem involves the contraction mapping principle. (Optional fun fact: The proof shows that the same kind of a result

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.
- The proof of this theorem involves the contraction mapping principle. (Optional fun fact: The proof shows that the same kind of a result holds when U, V are open subsets

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.
- The proof of this theorem involves the contraction mapping principle. (Optional fun fact: The proof shows that the same kind of a result holds when U, V are open subsets of a Banach space.

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.
- The proof of this theorem involves the contraction mapping principle. (Optional fun fact: The proof shows that the same kind of a result holds when U, V are open subsets of a Banach space.While we won't need this fact,

A special case - the inverse function theorem (IFT)

- Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.
- You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}.
- The proof of this theorem involves the contraction mapping principle. (Optional fun fact: The proof shows that the same kind of a result holds when U, V are open subsets of a Banach space.While we won't need this fact, it is useful for PDE.)

A consequence

A consequence

- Corollary:

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k
  (

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$.

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \;\forall\; a \in U$. Then F is an open map,

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \; \forall \; a \in U$. Then F is an open map, and if F is $1-1$,

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof:

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map:

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo.

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \; \forall \; a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set,

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \; \forall \; a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$.

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \; \forall \; a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$,

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof:

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \; \forall \; a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism:

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism: F is invertible.

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism: F is invertible. Thus, every local inverse coincide with

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism: F is invertible. Thus, every local inverse coincide with F^{-1} and hence F^{-1} is a C^k-diffeo. □

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism: F is invertible. Thus, every local inverse coincide with F^{-1} and hence F^{-1} is a C^k-diffeo. $\square$
- Therefore,

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism: F is invertible. Thus, every local inverse coincide with F^{-1} and hence F^{-1} is a C^k-diffeo. □
- Therefore, $x = r\cos(\theta), y = r\sin(\theta)$ is a

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \; \forall \; a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism: F is invertible. Thus, every local inverse coincide with F^{-1} and hence F^{-1} is a C^k-diffeo. □
- Therefore, $x = r\cos(\theta), y = r\sin(\theta)$ is a diffeo on $(0, \infty) \times (0, 2\pi)$ (

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \; \forall \; a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism: F is invertible. Thus, every local inverse coincide with F^{-1} and hence F^{-1} is a C^k-diffeo. □
- Therefore, $x = r\cos(\theta), y = r\sin(\theta)$ is a diffeo on $(0,\infty) \times (0, 2\pi)$ (more generally, on any open subset away from $r = 0$

A consequence

- Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k $(1 \leq k \leq \infty)$. Assume that $\det(DF_a) \neq 0 \; \forall \; a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
- Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
- Proof: Diffeomorphism: F is invertible. Thus, every local inverse coincide with F^{-1} and hence F^{-1} is a C^k-diffeo. □
- Therefore, $x = r\cos(\theta), y = r\sin(\theta)$ is a diffeo on $(0, \infty) \times (0, 2\pi)$ (more generally, on any open subset away from $r = 0$ where it is $1-1$).

Implicit function theorem (ImFT)

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y)

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k (

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$,

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof:

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT,

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT, G is a local C^k-diffeo from

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT, G is a local C^k-diffeo from $V_a \times W_b$ to $G(V_a \times W_b)$.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT, G is a local C^k-diffeo from $V_a \times W_b$ to $G(V_a \times W_b)$. $G^{-1}(x, c) = (x, y)$ iff $c = F(x, y)$.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT, G is a local C^k-diffeo from $V_a \times W_b$ to $G(V_a \times W_b)$. $G^{-1}(x, c) = (x, y)$ iff $c = F(x, y)$. Thus, $0 = F(x, y)$ iff $y = \pi_2 \circ G^{-1}(x, 0)$, i.e.,

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT, G is a local C^k-diffeo from $V_a \times W_b$ to $G(V_a \times W_b)$. $G^{-1}(x, c) = (x, y)$ iff $c = F(x, y)$. Thus, $0 = F(x, y)$ iff $y = \pi_2 \circ G^{-1}(x, 0)$, i.e., y is a C^k function of x locally.

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT, G is a local C^k-diffeo from $V_a \times W_b$ to $G(V_a \times W_b)$. $G^{-1}(x, c) = (x, y)$ iff $c = F(x, y)$. Thus, $0 = F(x, y)$ iff $y = \pi_2 \circ G^{-1}(x, 0)$, i.e., y is a C^k function of x locally. The derivative formula

Implicit function theorem (ImFT)

- Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k $(1 \leq k \leq \infty)$ function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.
- Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT, G is a local C^k-diffeo from $V_a \times W_b$ to $G(V_a \times W_b)$. $G^{-1}(x, c) = (x, y)$ iff $c = F(x, y)$. Thus, $0 = F(x, y)$ iff $y = \pi_2 \circ G^{-1}(x, 0)$, i.e., y is a C^k function of x locally. The derivative formula follows from the chain rule. □

A comment

A comment

- The proof shows that

A comment

- The proof shows that in fact,

A comment

- The proof shows that in fact, locally, $F(x,y) = c$ iff $y = G(x,c)$ where G is a C^k function. (

A comment

- The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense,

A comment

- The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense, we can "change coordinates"

A comment

- The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense, we can "change coordinates" in a C^k manner from

A comment

- The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense, we can "change coordinates" in a C^k manner from (x, y) to (x, c).

A comment

- The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense, we can "change coordinates" in a C^k manner from (x, y) to (x, c). In these new "coordinates",

A comment

- The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense, we can "change coordinates" in a C^k manner from (x, y) to (x, c). In these new "coordinates", the level set $F(x, y) = a$, looks like

A comment

- The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense, we can “change coordinates” in a C^k manner from (x, y) to (x, c). In these new “coordinates”, the level set $F(x, y) = a$, looks like $c = a$, i.e.,

A comment

- The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense, we can "change coordinates" in a C^k manner from (x, y) to (x, c). In these new "coordinates", the level set $F(x, y) = a$, looks like $c = a$, i.e., we "flatten" the level sets.)