

01: The Multivariate Chain Rule

Chapter Goal: To "upgrade" our understanding of the [Total Derivative](#) by formalizing its notation using the language of Linear Algebra, specifically vectors and the dot product.

1. Recap: The Total Derivative

$$f(x, y, z) = \sin(x) e^{yz^2}$$
$$x = t - 1; \quad y = t^2; \quad z = \frac{1}{t}$$
$$\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt}$$
$$\frac{df}{dt} = \cos(t - 1) e^{yz^2}$$

- **Recall (from Module 2):** If we have a function $f(x, y, z)$, where x , y , and z all depend on a single underlying variable t , then the total derivative $\frac{df}{dt}$ is:

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt}$$

- **Intuition:** The total change in f is the sum of the contributions from the change in each intermediate "path" (x , y , z).
-

2. Core Idea: Tidying Up the Notation with Linear Algebra

- **The Problem:** Writing out x, y, z, \dots and the long summation is tedious and inelegant, especially if we have thousands of variables.
- **The Solution:** Let's use the "language" of Linear Algebra to make it more concise.

Step 1: Pack the Inputs into a Vector

- Instead of writing $f(x_1, x_2, \dots, x_n)$, we will write it as $f(\vec{x})$.
- The bold symbol \vec{x} here is a **VECTOR** containing all the input variables:

$$\vec{x} = [x_1, x_2, \dots, x_n]^T.$$

Step 2: Identify the Two Groups of Derivatives

To build the Total Derivative, we need two main "ingredients":

- **"Ingredient A":** How does f change with respect to each x_i ?
 - This is the set of all partial derivatives of f .
 - Let's collect all these ingredients into a single vector. This is the **GRADIENT** of f .

$$\nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right]$$

- **"Ingredient B":** How does each x_i change with respect to t ?
 - This is the set of ordinary derivatives of each component of \vec{x} .
 - Let's also collect these into a vector.

$$\frac{d\vec{x}}{dt} = \begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{bmatrix}$$

3. "Aha!" Moment: The Connection to the Dot Product

Look again at the long formula for the Total Derivative:

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial f}{\partial x_2} \frac{dx_2}{dt} + \dots$$

What is the pattern of this calculation?

(Component₁ of A * Component₁ of B) + (Component₂ of A * Component₂ of B) + ...

This is the **DEFINITION** of the **[DOT PRODUCT](#)**!

$$f(\mathbf{x}) = f(x_1, x_2, x_3, \dots, x_n)$$

$$\frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \partial f / \partial x_3 \\ \vdots \\ \partial f / \partial x_n \end{bmatrix} \quad \frac{d\mathbf{x}}{dt} = \begin{bmatrix} dx_1/dt \\ dx_2/dt \\ dx_3/dt \\ \vdots \\ dx_n/dt \end{bmatrix} \quad \frac{df}{dt} = \begin{bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \partial f / \partial x_3 \\ \vdots \\ \partial f / \partial x_n \end{bmatrix} \cdot \begin{bmatrix} dx_1/dt \\ dx_2/dt \\ dx_3/dt \\ \vdots \\ dx_n/dt \end{bmatrix}$$

this is exactly what the dot product does.

4. The Elegant Final Formula: The Multivariate Chain Rule

By using vector notation, we can now rewrite the entire long Chain Rule into a single, beautiful, and concise expression:

$$\frac{df}{dt} = \nabla f \cdot \frac{d\vec{x}}{dt}$$

- **Translating the Formula:**

"The total rate of change of f with respect to t is equal to the **dot product** of the **Gradient of f** (how f changes with respect to \vec{x}) and the vector $\frac{d\vec{x}}{dt}$ (how \vec{x} changes with respect to t)."

- **Conclusion:**

- The **Multivariate Chain Rule** and the **Total Derivative** are two names for the exact same idea.
- By using Linear Algebra (vectors and the dot product), we can write this complex idea in a very compact and elegant way.
- This shows the power of the "language" of Linear Algebra to simplify problems in Calculus.

5. Concrete Analogy: Climbing a "Treasure Mountain"

Let's use the analogy of climbing a mountain again to make the geometric meaning of the formula $\frac{df}{dt} = \nabla f \cdot \frac{d\vec{x}}{dt}$ feel real.

Analogy Setup

- **The Landscape** $f(x, y)$: This is our mountain's elevation map. The peaks (high f values) have lots of gold.
- **You**: A hiker walking on this map.
- **Time** t : Your hiking time.

Two Important Vectors

1. Your Velocity Vector ($\frac{d\vec{x}}{dt}$):

- This is an arrow on the map describing your current movement.
- **Its Direction**: The compass direction you are currently walking (e.g., North-East).
- **Its Length**: How fast you are walking (e.g., 5 km/h).
- Let's call this $\vec{v} = \frac{d\vec{x}}{dt}$.

2. The Gradient Vector (∇f):

- This is your "treasure compass".
- It's an arrow on the map that points in the direction of the **steepest ascent** from your current position (e.g., straight North).

The Key Question

$\frac{df}{dt}$ is asking: "How fast is my gold elevation increasing right now?"

Analyzing the Formula: $\frac{df}{dt} = \nabla f \cdot \vec{v}$

The formula says that your rate of gold increase is the **dot product** between your "Treasure Compass" and your "Velocity Vector". Let's see what this means in three scenarios.

Scenario 1: The Most Efficient Movement

- **Situation**: You see your Treasure Compass (∇f) pointing straight North. Being very smart, you decide to walk exactly North (\vec{v}).
- **Dot Product Analysis**:
 - The vectors ∇f and \vec{v} are now perfectly parallel. The angle θ between them is 0° .
 - $\cos(0^\circ) = 1$.
 - $\frac{df}{dt} = \nabla f \cdot \vec{v} = |\nabla f||\vec{v}|\cos(0^\circ) = |\nabla f||\vec{v}|$.
- **Real Meaning**: Your rate of gold increase is at its **maximum**. You are getting 100% of the landscape's "steepness" multiplied by your walking speed. You are climbing in the most efficient way possible.

Scenario 2: A Wasted Movement

- **Situation:** The Treasure Compass (∇f) points North. But you decide to walk straight East (\vec{v}). Your direction is perpendicular to the direction of fastest ascent.
- **Dot Product Analysis:**
 - The vectors ∇f and \vec{v} are now mutually orthogonal. The angle θ is 90° .
 - $\cos(90^\circ) = 0$.
 - $\frac{df}{dt} = \nabla f \cdot \vec{v} = |\nabla f||\vec{v}| \cos(90^\circ) = 0$.
- **Real Meaning:** Your rate of gold increase is **zero**. You are walking along a **contour line** (a line of constant elevation). You are neither climbing nor descending. Your journey is wasted for the purpose of finding gold.

Scenario 3: A "Normal" Movement

- **Situation:** The Treasure Compass (∇f) points North. You are walking North-East (\vec{v}). Your direction is "somewhat" aligned with the ascent, but not perfectly.
- **Dot Product Analysis:**
 - The angle θ between "North" and "North-East" is 45° .
 - $\cos(45^\circ) \approx 0.707$.
 - $\frac{df}{dt} = \nabla f \cdot \vec{v} = |\nabla f||\vec{v}| \cos(45^\circ)$.
- **Real Meaning:** Your rate of gold increase is positive, but not maximal. You are only getting about **70.7%** of the maximum potential elevation gain. The $\cos(\theta)$ here acts as an **"efficiency factor"** for your direction of travel. You are climbing diagonally up the slope.

Conclusion

So, the formula $\frac{df}{dt} = \nabla f \cdot \frac{d\vec{x}}{dt}$ is not just a formula. It's a beautiful geometric statement:

"The total rate of change you experience ($\frac{df}{dt}$) is the projection of your movement velocity ($\frac{d\vec{x}}{dt}$) onto the direction of fastest change (∇f)."

It tells us that how fast we "climb" depends not only on how steep the mountain is ($|\nabla f|$) or how fast we are walking ($|\frac{d\vec{x}}{dt}|$), but also on how well our path is **aligned** with the ideal climbing direction. And that "alignment" is measured perfectly by the [Dot Product](#).

Note

Tentu saja. Sangat wajar jika ini terasa membingungkan. "Apa bedanya?" adalah pertanyaan yang sangat bagus, karena di permukaan, kelihatannya kita hanya mengubah cara menulisnya.

Mari kita bongkar ini dengan analogi yang sangat sederhana, tanpa matematika sama sekali.

Analogi: Menghitung Total Belanjaan

Bayangkan kamu sedang di kasir supermarket. Kamu membeli:

- 3 Apel
- 5 Jeruk
- 2 Pisang

Kasir perlu menghitung total biayanya.

Cara Lama (Cara "Anak SD" / Turunan Total Biasa)

Kasir akan menghitungnya satu per satu dan menjumlahkannya di kepalanya.

Proses Berpikir Kasir:

1. "Oke, apel dulu. 3 buah, harganya 2000 per buah. Jadi totalnya $3 * 2000 = 6000$."
2. "Sekarang jeruk. 5 buah, harganya 3000 per buah. Jadi totalnya $5 * 3000 = 15000$."
3. "Terakhir pisang. 2 buah, harganya 1500 per buah. Jadi totalnya $2 * 1500 = 3000$."
4. "Total semuanya adalah $6000 + 15000 + 3000 = 24000$."

Rumus Panjangnya:

$$\text{Total} = (\text{Jumlah_Apel} * \text{Harga_Apel}) + (\text{Jumlah_Jeruk} * \text{Harga_Jeruk}) + (\text{Jumlah_Pisang} * \text{Harga_Pisang})$$

Ini persis seperti rumus **Turunan Total** yang kita lihat sebelumnya:

$$df/dt = (\partial f/\partial x)*(dx/dt) + (\partial f/\partial y)*(dy/dt) + (\partial f/\partial z)*(dz/dt)$$

- Setiap suku $(...) * (...)$ dihitung secara terpisah, lalu semuanya dijumlahkan.
 - Cara ini **benar**, tapi **panjang** untuk ditulis dan **tidak terstruktur** dengan rapi.
-

Cara Baru (Cara "Manajer Gudang" / Aturan Rantai Multivariat)

Sekarang, bayangkan seorang manajer gudang yang lebih terorganisir. Dia tidak peduli dengan detail satu per satu. Dia berpikir dalam bentuk "**daftar**".

Proses Berpikir Manajer:

1. Buat Daftar #1 (Daftar Jumlah Barang):

- "Aku punya daftar jumlah barang yang dibeli."
- `Jumlah = [3, 5, 2]`

2. Buat Daftar #2 (Daftar Harga Barang):

- "Aku punya daftar harga untuk setiap jenis barang."
- `Harga = [2000, 3000, 1500]`

3. Lakukan Satu Operasi Tunggal:

- Manajer tidak perlu berpikir "apel... jeruk... pisang...". Dia hanya perlu melakukan satu operasi standar yang disebut **DOT PRODUCT** antara kedua daftar ini.
 - `Total = Jumlah · Harga`
 - `= [3, 5, 2] · [2000, 3000, 1500]`
 - `= (3*2000) + (5*3000) + (2*1500) = 24000`
-

Jadi, Apa Bedanya?

"apa bedanyaa dengan cara lamaa??"

TIDAK ADA bedanya dalam hal HASIL AKHIR. Perhitungan yang terjadi di balik layar sama persis.

Bedanya ada di CARA BERPIKIR dan CARA PENULISAN:

- **Cara Lama:** Berpikir secara **skalar**. Kamu melihat setiap "jalur" (x , y , z) sebagai entitas terpisah, menghitung kontribusinya satu per satu, lalu menjumlahkannya di akhir.
- **Cara Baru (Multivariat):** Berpikir secara **vektor**. Kamu **mengemas** semua informasi yang sejenis ke dalam "wadah" (vektor), lalu melakukan **satu operasi tunggal** (Dot Product) pada "wadah-wadah" itu.

Kenapa Cara Baru Lebih Baik?

1. **Ringkas:** $df/dt = \nabla f \cdot (dx/dt)$ jauh lebih pendek dan elegan untuk ditulis daripada $df/dt = (\partial f/\partial x_1)*(dx_1/dt) + \dots$ sampai n suku.
2. **Terstruktur:** Ini memisahkan masalah dengan rapi menjadi dua "benda":
 - ∇f : Semua informasi tentang bagaimana f berubah.
 - dx/dt : Semua informasi tentang bagaimana input x berubah.
3. **Siap untuk Komputer:** Bahasa vektor dan matriks adalah bahasa yang "dimengerti" oleh library komputasi seperti NumPy dan TensorFlow. Jauh lebih mudah untuk menulis kode `np.dot(grad_f, dxdt)` daripada menulis *loop* penjumlahan manual.

Kesimpulan:

Aturan Rantai Multivariat **bukanlah konsep baru**. Ia adalah **"bahasa" baru** yang lebih

canggih dan ringkas untuk mendeskripsikan Turunan Total yang sudah kita kenal, dengan meminjam kekuatan dari Aljabar Linear (vektor dan dot product).

⚠ Warning

Okee secara makna, arti di vektor ini ada artinyaa kahh?? atau bener bener cuman secara penulisan aja??? yang misalnyaa yang kek komputer kan gak jago kan yaa kalau misalnyaaa kita nulisnyaa kek

$$3a + 5b + 7c = 10$$

$$2a + 4b + 3c = 15$$

.....

dia mending ditulis dalam bentuk matrikss yaa kan?? nah apakah ini itu bener bener sekedar kek cara kita ngomong ke komputer biar dia paham dan lebih mudahh bagi komputer, meskipun secara pengerjaan, hasil akhir dan sebagainya samaa??

Jawaban singkatnya: **Ini dimulai sebagai penyederhanaan penulisan, tapi ternyata membuka pintu ke pemahaman geometris yang jauh lebih dalam.**

Level 1: Penyederhanaan Notasi (Seperti yang Kamu Bilang)

"nah apakah ini itu bener bener sekedar kek cara kita ngomong ke komputer biar dia paham dan lebih mudahh bagi komputer, meskipun secara pengerjaan, hasil akhir dan sebagainya samaa??"

YA. Pada level yang paling dasar, kamu 100% benar.

$$df/dt = (\partial f/\partial x)*(dx/dt) + (\partial f/\partial y)*(dy/dt) + (\partial f/\partial z)*(dz/dt)$$

vs.

$$df/dt = \nabla f \cdot (dx/dt)$$

Kedua baris di atas secara matematis **IDENTIK**. Komputer pada akhirnya akan melakukan operasi "kali-lalu-jumlah" yang sama persis.

Menulisnya dalam bentuk Dot Product itu seperti:

- **Menggunakan Fungsi SUM() di Excel:** Daripada menulis `=A1+A2+A3+...+A100`, kamu menulis `=SUM(A1:A100)`. Hasilnya sama, tapi cara menulisnya jauh lebih bersih, efisien, dan mengurangi kemungkinan salah ketik.

- **Menggunakan Notasi Matriks untuk Sistem Persamaan:** Kamu benar sekali di sini. Menulis $Ax=v$ jauh lebih ringkas daripada menulis 100 persamaan linear.

Jadi, pada level pertama, ini adalah kemenangan besar untuk **keringkasan dan kejelasan notasi**. Ini membantu kita "berbicara" tentang ide-ide yang kompleks dengan lebih mudah.

Level 2: Munculnya Makna Geometris (Di Sinilah Keajaibannya)

Di sinilah jawabannya lebih dari sekadar "penulisan". Begitu kita mengemasnya ke dalam bentuk vektor, kita secara tidak sengaja **memberikan makna geometris baru** pada setiap bagian.

Mari kita lihat lagi:

$$df/dt = \nabla f \cdot (dx/dt)$$

Sekarang ∇f dan dx/dt bukan lagi sekadar "daftar angka". Kita bisa memikirkannya sebagai **PANAH** di ruang multi-dimensi.

- **Vektor dx/dt** : Ini adalah "**Vektor Kecepatan**". Panah ini memberitahu kita **ke arah mana** dan **seberapa cepat** kita sedang bergerak di "ruang input" (x, y, z).
- **Vektor ∇f** : Ini adalah "**Vektor Gradien**". Panah ini memberitahu kita **arah tanjakan tercepat** dan **seberapa curam** "lanskap" fungsi f di posisi kita saat ini.

Sekarang, lihat kembali operasinya: **DOT PRODUCT**.

Apa arti geometris dari Dot Product?

$A \cdot B = |A| |B| \cos(\theta)$. Ini adalah tentang **PROYEKSI**. Dot product mengukur "**seberapa sejajar**" dua vektor itu.

"Aha!" Moment Geometris:

Sekarang kita bisa "membaca" Aturan Rantai dengan cara yang sama sekali baru:

$$df/dt = \nabla f \cdot (dx/dt)$$

"Laju perubahan total (df/dt)...

...adalah hasil dari **memproyeksikan** 'Vektor Kecepatan' kita (dx/dt)...

...ke 'Vektor Arah Tanjakan Tercepat' (∇f)."

Artinya secara Intuitif:

Perubahan total yang kamu rasakan (df/dt) akan **maksimal** jika kamu bergerak **persis searah** dengan Gradien (sejajar dengan arah tanjakan tercepat).

Perubahan totalnya akan **nol** jika kamu bergerak **tegak lurus** terhadap Gradien (berjalan di sepanjang garis kontur).

Perubahan totalnya akan **negatif** jika kamu bergerak **berlawanan arah** dengan Gradien (menuruni lembah).

Kesimpulan:

- Apakah ini hanya penyederhanaan penulisan? **Ya, pada awalnya.**
- Apakah ada artinya? **YA, SANGAT!** Dengan menulisnya sebagai dot product, kita membuka pintu untuk **interpretasi geometris** yang sangat kuat. Kita tidak lagi melihatnya sebagai penjumlahan mekanis, tapi sebagai interaksi antara dua "panah": **panah gerakan kita** dan **panah kecuraman lanskap**.

Ini adalah pola yang berulang di matematika: notasi yang baik tidak hanya membuat sesuatu lebih ringkas, tapi juga seringkali mengungkap struktur dan makna yang lebih dalam yang sebelumnya tersembunyi.

Tags: [#mml-specialization](#) [#multivariate-calculus](#) [#chain-rule](#) [#total-derivative](#) [#gradient](#)