FuVar Notes

Naoki Pross – naoki.pross@ost.ch

Spring Semseter 2021

1 Preface

These are just my personal notes of the FuVar course, and definitively not a rigorously constructed mathematical text. The good looking LaTeX typesetting may trick you into thinking it is rigorous, but really, it is not.

2 Derivatives of vector valued functions

Definition 1 (Partial derivative)**.** A vector values function $f : \mathbb{R}^m \to \mathbb{R}$, with $\mathbf{v} \in \mathbb{R}^m$, has a partial derivative with respect to v_i defined as

$$\partial_{v_i} f(\mathbf{v}) = f_{v_i}(\mathbf{v}) = \lim_{h \to 0} \frac{f(\mathbf{v} + h\mathbf{e}_j) - f(\mathbf{v})}{h}$$

Theorem 1. (Schwarz's theorem, symmetry of partial derivatives) Under some generally satisfied conditions (continuity of n-th order partial derivatives) Schwarz's theorem states that it is possible to swap the order of differentiation.

$$\partial_x \partial_y f(x, y) = \partial_y \partial_x f(x, y)$$

Definition 2 (Linearization)**.** A function $f : \mathbb{R}^m \to \mathbb{R}$ has a linearization g at $\mathbf{x}_0$ given by

$$g(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{i=1}^{m} \partial_{x_i} f(\mathbf{x}_0)(x_i - x_{i,0}),$$

if all partial derviatives are defined at $\mathbf{x}_0$.

Theorem 2 (Propagation of uncertanty)**.** Given a measurement of m values in a vector $\mathbf{x} \in \mathbb{R}^m$ with values given in the form $x_i = \bar{x}_i \pm \sigma_{x_i}$, a linear approximation the error of a dependent variable y is computed with

$$y = \bar{y} \pm \sigma_y \approx f(\bar{\mathbf{x}}) \pm \sqrt{\sum_{i=1}^{m} \left(\partial_{x_i} f(\bar{\mathbf{x}}) \sigma_{x_i}\right)^2}$$

Definition 3 (Gradient vector)**.** The *gradient* of a function $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^m$ is a column vector[1] containing the derivatives in each direction.

$$\boldsymbol{\nabla} f(\mathbf{x}) = \sum_{i=1}^{m} \partial_{x_i} f(\mathbf{x}) \mathbf{e}_i = \begin{pmatrix} \partial_{x_1} f(\mathbf{x}) \\ \vdots \\ \partial_{x_m} f(\mathbf{x}) \end{pmatrix}$$

Definition 4 (Directional derivative)**.** A function $f(\mathbf{x})$ has a directional derivative in direction $\mathbf{r}$ (with $|\mathbf{r}| = 1$) given by

$$\frac{\partial f}{\partial \mathbf{r}} = \nabla \mathbf{r} f = \mathbf{r} \cdot \boldsymbol{\nabla} f$$

Theorem 3. The gradient vector always points towards *the direction of steepest ascent*.

Definition 5 (Jacobian Matrix)**.** The *Jacobian* $\mathbf{J}_f$ (sometimes written as $\frac{\partial(f_1, \dots f_m)}{\partial(x_1, \dots, x_n)}$) of a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix $\in \mathbb{R}^{n \times m}$ whose entry at the i-th row and j-th column is given by $(\mathbf{J}_f)_{i,j} = \partial_{x_j} f_i$, so

$$\mathbf{J}_f = \begin{pmatrix} \partial_{x_1} f_1 & \cdots & \partial_{x_n} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \cdots & \partial_{x_n} f_m \end{pmatrix} = \begin{pmatrix} (\boldsymbol{\nabla} f_1)^t \\ \vdots \\ (\boldsymbol{\nabla} f_m)^t \end{pmatrix}$$

Remark 1. In the scalar case ($m = 1$) the Jacobian matrix is the transpose of the gradient vector.

Definition 6 (Hessian matrix)**.** Given a function $f : \mathbb{R}^m \to \mathbb{R}$, the square matrix whose entry at the i-th row and j-th column is the second derivative of f first with respect to x_j and then to x_i is know as the *Hessian* matrix. $(\mathbf{H}_f)_{i,j} = \partial_{x_i} \partial_{x_j} f$ or

$$\mathbf{H}_f = \begin{pmatrix} \partial_{x_1} \partial_{x_1} f & \cdots & \partial_{x_1} \partial_{x_m} f \\ \vdots & \ddots & \vdots \\ \partial_{x_m} \partial_{x_1} f & \cdots & \partial_{x_m} \partial_{x_m} f \end{pmatrix}$$

Because (almost always) the order of differentiation does not matter, it is a symmetric matrix.

[1] In matrix notation it is also often defined as row vector to avoid having to do some transpositions in the Jacobian matrix and dot products in directional derivatives

3 Methods for maximization and minimization problems

Method 1 (Find stationary points)**.** Given a function $f : D \subseteq \mathbb{R}^m \to \mathbb{R}$, to find its maxima and minima we shall consider the points

- that are on the boundary of the domain ∂D,
- where the gradient $\boldsymbol{\nabla} f$ is not defined,
- that are stationary, i.e. where $\boldsymbol{\nabla} f = \mathbf{0}$.

Method 2 (Determine the type of stationary point for 2 dimensions)**.** Given a scalar function of two variables $f(x, y)$ and a stationary point $\mathbf{x}_s$ (where $\boldsymbol{\nabla} f(\mathbf{x}_s) = \mathbf{0}$), we define the *discriminant*

$$\Delta = \partial_x^2 f \partial_y^2 f - \partial_y \partial_x f$$

- if $\Delta > 0$ then $\mathbf{x}_s$ is an extrema, if $\partial_x^2 f(\mathbf{x}_s) < 0$ it is a maximum, whereas if $\partial_x^2 f(\mathbf{x}_s) > 0$ it is a minimum;
- if $\Delta < 0$ then $\mathbf{x}_s$ is a saddle point;
- if $\Delta = 0$ we need to analyze further.

Remark 2. The previous method is obtained by studying the second directional derivative $\nabla_{\mathbf{r}} \nabla_{\mathbf{r}} f$ at the stationary point in direction of a vector $\mathbf{r} = \mathbf{e}_1 \cos(\alpha) + \mathbf{e}_2 \sin(\alpha)$

Method 3 (Determine the type of stationary point in higher dimensions)**.** Given a scalar function of two variables $f(x, y)$ and a stationary point $\mathbf{x}_s$ (where $\boldsymbol{\nabla} f(\mathbf{x}_s) = \mathbf{0}$), we compute the Hessian matrix $\mathbf{H}_f(\mathbf{x}_s)$. Then we compute its eigenvalues $\lambda_1, \ldots, \lambda_m$ and

- if all $\lambda_i > 0$, the point is a minimum;
- if all $\lambda_i < 0$, the point is a maximum;
- if there are both positive and negative eigenvalues, it is a saddle point.

In the other cases, when there are $\lambda_i \leq 0$ and/or $\lambda_i \geq 0$ further analysis is required.

Remark 3. Recall that to compute the eigenvalues of a matrix, one must solve the equation $(\mathbf{H} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$. Which can be done by solving the characteristic polynomial $\det(\mathbf{H} - \lambda \mathbf{I}) = 0$ to obtain $\dim(\mathbf{H})$ λ_i, which when plugged back in result in a overdetermined system of equations.

Method 4 (Quickly find the eigenvalues of a 2×2 matrix)**.** Let

$$m = \frac{1}{2} \operatorname{tr} \mathbf{H} = \frac{a + d}{2}, \qquad p = \det \mathbf{H} = ad - bc,$$

then $\lambda = m \pm \sqrt{m^2 - p}$.

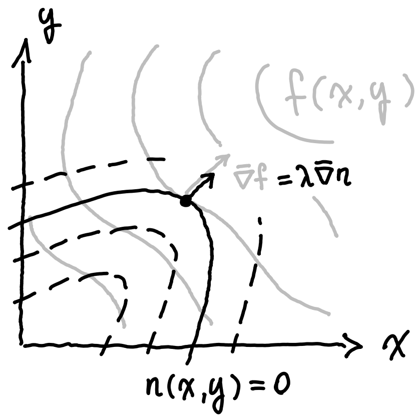

Figure 1: Intuition for the method of Lagrange multipliers. Extrema of a constrained function are where $\boldsymbol{\nabla} f$ is proportional to $\boldsymbol{\nabla} n$.

Method 5 (Search for a constrained extremum in 2 dimensions)**.** Let $n(x, y) = 0$ be a constraint in the search of the extrema of a function $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$. To find the extrema we look for points

- on the boundary $\mathbf{u} \in \partial D$ where $n(\mathbf{u}) = 0$;
- $\mathbf{u}$ where the gradient either does not exist or is $\mathbf{0}$, and satisfy $n(\mathbf{u}) = 0$;
- that solve the system of equations

$$\begin{cases} \partial_x f(\mathbf{u}) \cdot \partial_y n(\mathbf{u}) = \partial_y f(\mathbf{u}) \cdot \partial_x n(\mathbf{u}) \\ n(\mathbf{u}) = 0 \end{cases}$$

Method 6 (Search for a constrained extremum in higher dimensions, method of Lagrange multipliers)**.** We wish to find the extrema of $f : D \subseteq \mathbb{R}^m \to \mathbb{R}$ under $k < m$ constraints $n_1 = 0, \cdots, n_k = 0$. To find the extrema we consider the following points:

- Points on the boundary $\mathbf{u} \in \partial D$ that satisfy $n_i(\mathbf{u}) = 0$ for all $1 \leq i \leq k$,
- Points $\mathbf{u} \in D$ where either
  - any of $\boldsymbol{\nabla} f, \boldsymbol{\nabla} n_1, \ldots, \boldsymbol{\nabla} n_k$ do not exist, or
  - $\boldsymbol{\nabla} n_1, \ldots, \boldsymbol{\nabla} n_k$ are linearly *dependent*,

  and that satisfy $0 = n_1(\mathbf{u}) = \ldots = n_k(\mathbf{u})$.
- Points that solve the system of $m + k$ equations

$$\begin{cases} \boldsymbol{\nabla} f(\mathbf{u}) = \sum_{i=1}^{k} \lambda_i \boldsymbol{\nabla} n_i(\mathbf{u}) & (m\text{-dimensional}) \\ n_i(\mathbf{u}) = 0 & \text{for } 1 \leq i \leq k \end{cases}$$

  The λ values are known as *Lagrange multipliers*. The same calculation can be written more

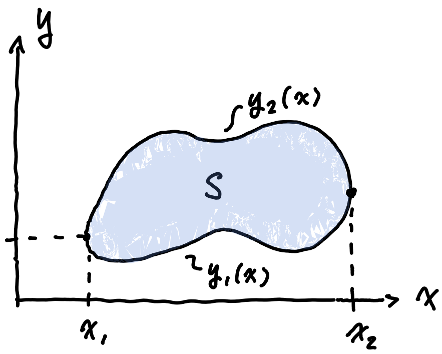

Figure 2: Double integral.

compactly by defining the $m+k$ dimensional *Lagrangian*

$$\mathcal{L}(\mathbf{u}, \boldsymbol{\lambda}) = f(\mathbf{u}) - \sum_{i=0}^{k} \lambda_i n_i(\mathbf{u})$$

where $\boldsymbol{\lambda} = \lambda_1, \ldots, \lambda_k$ and then solving $\boldsymbol{\nabla}\mathcal{L}(\mathbf{u}, \boldsymbol{\lambda}) = \mathbf{0}$. This is generally used in numerical computations and not very useful by hand.

4 Integration of vector values scalar functions

Theorem 4 (Change the order of integration for double integrals)**.** For a double integral over a region S (see Fig. 2) we need to compute

$$\iint_S f(x,y)\, ds = \int\limits_{x_1}^{x_2} \int\limits_{y_1(x)}^{y_2(x)} f(x,y)\, dydx.$$

If $y_1(x)$ and $y_2(x)$ are bijective we can swap the order of integration by finding the inverse functions $x_1(y)$ and $x_2(y)$. If they are not bijective (like in Fig. 2), the region must be split into smaller parts. If the region is a rectangle it is always possible to change the order of integration.

Theorem 5 (Transformation of coordinates in 2 dimensions)**.** Given two "nice" functions $x(u,v)$ and $y(u,v)$, that means are a bijection from S to S' with continuous partial derivatives and nonzero Jacobian determinant $|\mathbf{J}_f| = \partial_u x \partial_v y - \partial_v x \partial_u y$, which transform the coordinate system. Then

$$\iint_S f(x,y)\, ds = \iint_{S'} f(x(u,v), y(u,v)) |\mathbf{J}_f|\, ds$$

Theorem 6 (Transformation of coordinates)**.** The generalization of theorem 5 is quite simple. For an n-integral of a function $f : \mathbb{R}^m \to \mathbb{R}$ over a region B, we let $\mathbf{x}(\mathbf{u})$ be "nice" functions that transform the coordinate system. Then as before

$$\int_B f(\mathbf{x})\, ds = \int_{B'} f(\mathbf{x}(\mathbf{u})) |\mathbf{J}_f|\, ds$$

	Volume dv	Surface ds
Cartesian	–	$dx\, dy$
Polar	–	$r d\, r d\phi$
Curvilinear	–	$\lvert\mathbf{J}_f\rvert\, du\, dv$
Cartesian	$dx\, dy\, dz$	$\hat{\mathbf{z}}\, dx\, dy$
Cylindrical	$r\, dr\, d\phi\, dz$	$\hat{\mathbf{z}} r\, dr\, d\phi$
		$\hat{\boldsymbol{\phi}}\, dr\, dz$
		$\hat{\mathbf{r}} r\, d\phi\, dz$
Spherical	$r^2 \sin\theta\, dr\, d\theta\, d\phi$	$\hat{\mathbf{r}} r^2 \sin\theta\, d\theta\, d\phi$
Curvilinear	$\lvert\mathbf{J}_f\rvert\, du\, dv\, dw$	–

Table 1: Differential elements for integration.

5 Derivatives of curves

License