Introduction to Optimization with Applications to PDE Numerical Algorithms
Instructors: Faker BELGACEM[1], Marie POSTEL[2]
Memory help

Basic Linear Algebra Formulas

Vector Operations

- Addition: $\mathbf{v} + \mathbf{w} = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$
- Scalar Multiplication: $c \cdot \mathbf{v} = c\mathbf{v} = (c \cdot v_1, c \cdot v_2, \ldots, c \cdot v_n) = (cv_1, cv_2, \ldots, cv_n)$
- Dot Product: $\mathbf{v} \cdot \mathbf{w} = v_1 \cdot w_1 + v_2 \cdot w_2 + \ldots + v_n \cdot w_n = \sum_{i=1}^n v_i w_i$

Matrix Operations

- Addition: $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij}])_{i,j}$
- Scalar Multiplication: $c \cdot \mathbf{A} = (c \cdot a_{ij}) = (ca_{ij})$
- Transpose: $(\mathbf{A}^T)_{ij} = a_{ji}$
- Matrix Multiplication: $\mathbf{C} = \mathbf{A} \cdot \mathbf{B} = \mathbf{AB}$, where $c_{ij} = \sum_{k=1}^n a_{ik} \cdot b_{kj} = \sum_{k=1}^n a_{ik} b_{kj}$
- $\mathbf{v} \cdot \mathbf{A}\mathbf{w} = \mathbf{A}^\mathbf{T}\mathbf{v} \cdot \mathbf{w}$

Matrix results

- Matrix ensembles

$$\begin{aligned} S_n &= \{A \in \mathcal{M}_{n\times n}(\mathbb{R}),\ a_{i,j} = a_{j,i}\}, \\ S_n^+ &= \{A \in S_n, Av \cdot v \geq 0\ \forall v \in \mathbb{R}^n\}, \\ S_n^{++} &= \{A \in S_n^+, Av \cdot v > 0\ \forall v \in \mathbb{R}^n \backslash \{0\}\} \end{aligned}$$

- Spectrum $Sp(A) = \{\lambda \in \mathbb{R},\ \exists v \neq 0_{\mathbb{R}^n}, Av = \lambda v\}$, eigenvalues λ, eigenvectors v
- Matrix norm subordinate to a vector norm $\|A\| = \max_{v\in\mathbb{R}^n} \dfrac{\|Av\|}{\|v\|} = \max_{\|v\|=1} \|Av\|$
- Frobenius norm : $\|A\| = \mathrm{tr}(A^T A) = \sqrt{\sum_{i,j} a_{ij}^2}$
- Euclidian norm : If $\|v\|^2 = \sum_{i=1}^n v_i^2$ then $\|A\| = \sqrt{\rho(A)}$ with $\rho(A) = \max_{Sp(A^T A)} \lambda$
- If $A \in S_n$ then
  - $Sp(A) \subset \mathbb{R} : \lambda_{\min} = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n = \lambda_{\max}$
  - $\lambda_{\min}\|v\|^2 \leq Av \cdot v \leq \lambda_{\max}\|v\|^2$, $\forall v \in \mathbb{R}^n$
  - $\|A\| = \max_{\lambda\in Sp(A)} |\lambda|$

Basic Differential Calculus

- E, F, Banach spaces,
  $f : E \to F$, $f(h) = o(h^p)$ if and only if $\lim_{h\to 0} \dfrac{f(h)}{h^p} = 0$.
  $f(h) = O(h^p)$ if and only if $\exists C \geq 0$, $\exists \varepsilon > 0$, $\|h\| \leq \varepsilon \Rightarrow \|f(h)\| \leq C\|h\|^p$.
- $f : E \to F$ differentiable at $x \in E$ if and only if

$$\exists Df(x) \in \mathcal{L}(E, F), \quad \forall h \in E, \quad f(x + h) = f(x) + Df(x)(h) + o(\|h\|.$$

- $f : \mathbb{R}^n \to \mathbb{R}$ differentiable at $x \in \mathbb{R}^n$. We denote $\nabla f(x)$ the column vector gradient $\left(\dfrac{\partial f}{\partial x_i}\right)_{i=1,\ldots,n}$
  - $\forall h \in \mathbb{R}^n$

$$\begin{aligned} f(x + h) &= f(x) + Df(x)(h) + o(\|h\|), \\ &= f(x) + \nabla f(x) \cdot h + o(\|h\|) \end{aligned}$$

[1]Laboratoire de Mathématiques Appliquées de Compiègne
[2]Laboratoire Jacques-Louis Lions, Sorbonne Université

- $\forall h \in \mathbb{R}^n$, $\exists \eta \in]0,1[$

$$f(x+h) \quad = \quad f(x) + \nabla f(x+\eta h) \cdot h.$$

- $f : \mathbb{R}^n \to \mathbb{R}^m$ differentiable at $x \in \mathbb{R}^n$, $f(x) = (f_i(x))_{i=1,\cdots,m}$, $f_i : \mathbb{R}^n \to \mathbb{R}$.
  We denote $Jf(x) \in \mathcal{M}_{m\times n}(\mathbb{R})$ the Jacobian matrix $Jf(x)_{i,j} = \dfrac{\partial f_{i,j}(x)}{\partial x_j}$ for $i = 1,\dots,m$, $j = 1,\dots,n$
- $f : \mathbb{R}^n \to \mathbb{R}$ twice differentiable at $x \in \mathbb{R}^n$. We denote $Hf(x)$ -sometimes $\nabla^2 f(x)$)- the Hessian matrix $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\dots,n}$
  - $\forall h \in \mathbb{R}^n$

$$f(x+h) \quad = \quad f(x) + \nabla f(x) \cdot h + \frac{1}{2} Hf(x)h \cdot h + o(\|h\|^2)$$

  - $\forall h \in \mathbb{R}^n$, $\exists \eta \in]0,1[$

$$f(x+h) \quad = \quad f(x) + \nabla f(x) \cdot h + \frac{1}{2} Hf(x+\eta h)h \cdot h$$

Convexity

- Sets: a set C is convex if it contains all the line segments $\{x = \theta x_1 + (1-\theta)x_2, \quad 0 \leq \theta \leq 1\}$ connecting any two points of C.
  Examples
  - Ball (Euclidean) with center x_c and radius r
    $B(x_c, r) = \{x, \|x - x_c\|_2 \leq r\} = \{x_c + ru, \|u\|_2 \leq 1\}$
  - Ellipsoid
    $E = \{x, (x-x_c)^t P^{-1}(x-x_c) \leq 1\}$ with $P \in S^n_{++}$
  - Polyhedron: set of the form $P := \{x \in \mathbb{R}^n, Ax \preceq b\}$ with $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ fixed.
- Functions
  - A function $f : C \longrightarrow \mathbb{R}$ is
    * convex iff for all $x, y \in C$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y), \quad \forall \lambda \in [0,1].$$

    * strictly convex iff for all $x, y \in C$, $x \neq y$

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y), \quad \forall \lambda \in]0,1[.$$

  - Examples on $\mathbb{R}^n$
    * Affine $f(x) = a^t x + b$, for $a, x \in \mathbb{R}^n$, $b \in \mathbb{R}$
    * Norms $\|x\|_p = (\sum_{i=1}^n x_i^p)^{1/p}$ for $p \geq 1$, $\|x\|_\infty = \max_k |x_k|$,
  - Examples on $\mathcal{M}_{m\times n}(\mathbb{R})$
    * $f(X) = tr(A^T X) + b = \sum_{i=1}^m \sum_{j=1}^n A_{ij} X_{ij} + b$
    * Spectral Norm: Maximum Singular Value $f(X) = \|X\|_2 = \sigma_{\max}(X) = (\lambda_{\max}(X^T X))^{1/2}$
- $f : \mathbb{R}^n \to \mathbb{R}$. We denote $g_{x,v}(t) = f(x+tv)$ with $x, v \in \mathbb{R}^n$, $t \in \mathbb{R}$
  f is convex iff $g_{x,v}(t)$ convex for all $x, v \in \mathbb{R}^n$
- Let f be a function defined on a real-valued differentiable convex $C \subset \mathbb{R}^n$. The function f is
  - convex if and only if $\forall (x,y) \in C^2, \langle \nabla f(x), y - x\rangle \leq f(y) - f(x)$.
  - convex if and only if $\forall (x,y) \in C^2, \langle \nabla f(x) - \nabla f(y), x - y\rangle \geq 0$.
  - strictly convex if and only if $\forall (x,y) \in C^2, x \neq y, \langle \nabla f(x), y - x\rangle < f(y) - f(x)$.

  $f : \mathbb{R}^n \to \mathbb{R}$ twice differentiable on $\mathrm{dom} f$ convex, Hessian matrix $Hf(x) \in S^n$
  - f is convex iff $Hf(x) \in S^n_+ \quad \forall x \in \mathrm{dom} f$ (positive semi-definite)
  - if $Hf(x) \in \ S^n_{++} \quad \forall x \in \mathrm{dom} f$ (positive definite) then f is strictly convex