Math 516: Numerical Optimization

Lecture based on

Convex Analysis and Nonsmooth Optimization

by Dmitriy Drusvyatskiy

Background Material

Inner Products

Throughout, $\mathbf{E}$ is a *Euclidean space* ,
i.e., a finite-dim real vector space with an *inner product* $\langle \cdot, \cdot \rangle$.
Occasionally we say that $(\mathbf{E}, \langle \cdot, \cdot \rangle)$ is the Euclidean space when the choice of inner product needs to be specified.

Recall that an inner-product on $\mathbf{E}$ is an assignment
$\langle \cdot, \cdot \rangle \colon \mathbf{E} \times \mathbf{E} \to \mathbf{R}$ satisfying the following three properties for all $x, y, z \in \mathbf{E}$ and scalars $a, b \in \mathbf{R}$:

(Symmetry) $\langle x, y \rangle = \langle y, x \rangle$

(Bilinearity) $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$

(Positive definiteness) $\langle x, x \rangle \geq 0$ and equality $\langle x, x \rangle = 0$ holds if and only if $x = 0$.

Examples of Inner Products

Standard ip for $\mathbf{R}^n$: $\langle x, y \rangle := \sum_{i=1}^n x_i y_i = \|x\| \, \|y\| \cos\theta$, where θ is the angle between x and y.

Standard ip for $\mathbf{R}^{m\times n}$: The Frobenius or trace inner product,

$$\langle X, Y \rangle := \operatorname{tr} X^T Y = \sum_{i,j} X_{ij} Y_{ij}.$$

Real polynomials in one variable of degree $\leq n$ on $[a, b]$:

Integration inner product

$$\langle p, q \rangle := \int_a^b p(t) q(t) dt.$$

Adjoints of Linear Transformations

Suppose both $(\mathbf{X}, \langle\cdot,\cdot\rangle_{\mathbf{X}})$ and $(\mathbf{Y}, \langle\cdot,\cdot\rangle_{\mathbf{Y}})$ are Euclidean spaces.

Let $\mathcal{A} \in \mathbf{L}(\mathbf{X}, \mathbf{Y})$ where $\mathbf{L}(\mathbf{X}, \mathbf{Y})$ is the vector space of linear operators (or linear transformations) from $\mathbf{X}$ to $\mathbf{Y}$.

There exists a unique linear mapping $\mathcal{A}^*\colon \mathbf{Y} \to \mathbf{X}$, called the *adjoint*, satisfying

$$\langle \mathcal{A}^* y, x\rangle_{\mathbf{X}} = \langle y, \mathcal{A}x\rangle_{\mathbf{Y}} \qquad \text{for all points} \qquad x \in \mathbf{X},\, y \in \mathbf{Y}.$$

When $\mathbf{X} = \mathbf{R}^n$ and $\mathbf{Y} = \mathbf{R}^m$, every linear map $\mathcal{A}$ can be identified with a matrix $A \in \mathbf{R}^{m\times n}$. In this case, the matrix associated with the adjoint $\mathcal{A}^*$ is the transpose A^T.

Note: The adjoint differs significantly from the *classical adjoint* in Cramer's Rule.

Self-adjoint Linear Operators

Let $(\mathbf{E}, \langle\cdot,\cdot\rangle)$ be a Euclidean space and let $\mathcal{A} \in \mathbf{L}(\mathbf{E}, \mathbf{E})$.

We say that $\mathcal{A}$ is *self-adjoint* if $\mathcal{A} = \mathcal{A}^*$. The set of all self-adjoint linear operators on $\mathbf{E}$ is denoted by $\mathcal{S}(\mathbf{E})$ or $\mathcal{S}(\mathbf{E}, \langle\cdot,\cdot\rangle)$ if great specificity is required.

If $\mathbf{E} = \mathbf{R}^n$, the matrix representation of a self-adjoint linear operator is a symmetric matrix.

A self-adjoint linear operator on $\mathbf{R}^n$ can be identified with the symmetric matrices on $\mathbf{R}^n$ and so form a subspace of $\mathbf{R}^{n\times n}$ which we denote by $\mathbf{S}^n := \{A \in \mathbf{R}^{n\times n} \mid A = A^T\}$.

Self-adjoint Linear Operators

Let $(\mathbf{E}, \langle\cdot,\cdot\rangle)$ be a Euclidean space and let $\mathcal{A} \in \mathbf{L}(\mathbf{E}, \mathbf{E})$.

We say that $\mathcal{A}$ is *self-adjoint* if $\mathcal{A} = \mathcal{A}^*$. The set of all self-adjoint linear operators on $\mathbf{E}$ is denoted by $\mathcal{S}(\mathbf{E})$ or $\mathcal{S}(\mathbf{E}, \langle\cdot,\cdot\rangle)$ if great specificity is required.

If $\mathbf{E} = \mathbf{R}^n$, the matrix representation of a self-adjoint linear operator is a symmetric matrix.

A self-adjoint linear operator on $\mathbf{R}^n$ can be identified with the symmetric matrices on $\mathbf{R}^n$ and so form a subspace of $\mathbf{R}^{n\times n}$ which we denote by $\mathbf{S}^n := \{A \in \mathbf{R}^{n\times n} \,|\, A = A^T\}$.

Example: Let $A \in \mathbf{S}^n$ and define $\mathcal{H} : \mathbf{S}^n \to \mathbf{S}^n$ by $\mathcal{H}(X) := AXA$. Then $\mathcal{H} \in \mathcal{S}(\mathbf{S}^n)$ is a self-adjoint linear operator.

Self-adjoint Linear Operators

Let $(\mathbf{E}, \langle \cdot, \cdot \rangle)$ be a Euclidean space and let $\mathcal{A} \in \mathbf{L}(\mathbf{E}, \mathbf{E})$.

We say that $\mathcal{A}$ is *self-adjoint* if $\mathcal{A} = \mathcal{A}^*$. The set of all self-adjoint linear operators on $\mathbf{E}$ is denoted by $\mathcal{S}(\mathbf{E})$ or $\mathcal{S}(\mathbf{E}, \langle \cdot, \cdot \rangle)$ if great specificity is required.

If $\mathbf{E} = \mathbf{R}^n$, the matrix representation of a self-adjoint linear operator is a symmetric matrix.

A self-adjoint linear operator on $\mathbf{R}^n$ can be identified with the symmetric matrices on $\mathbf{R}^n$ and so form a subspace of $\mathbf{R}^{n \times n}$ which we denote by $\mathbf{S}^n := \{A \in \mathbf{R}^{n \times n} \mid A = A^T\}$.

Example: Let $A \in \mathbf{S}^n$ and define $\mathcal{H} : \mathbf{S}^n \to \mathbf{S}^n$ by $\mathcal{H}(X) := AXA$. Then $\mathcal{H} \in \mathcal{S}(\mathbf{S}^n)$ is a self-adjoint linear operator. How do we obtain a matrix representation for $\mathcal{H}$?

Positive Semi-Definite Linear Operators

A self-adjoint operator $\mathcal{A}$ is *positive semi-definite*, denoted $\mathcal{A} \succeq 0$, whenever

$$\langle \mathcal{A}x, x \rangle \geq 0 \quad \text{for all } x \in \mathbf{E}.$$

Similarly, a self-adjoint operator $\mathcal{A}$ is *positive definite*, denoted $\mathcal{A} \succ 0$, whenever

$$\langle \mathcal{A}x, x \rangle > 0 \quad \text{for all } 0 \neq x \in \mathbf{E}.$$

Positive Semi-Definite Linear Operators

A self-adjoint operator $\mathcal{A}$ is *positive semi-definite*, denoted $\mathcal{A} \succeq 0$, whenever

$$\langle \mathcal{A}x, x \rangle \geq 0 \quad \text{for all } x \in \mathbf{E}.$$

Similarly, a self-adjoint operator $\mathcal{A}$ is *positive definite*, denoted $\mathcal{A} \succ 0$, whenever

$$\langle \mathcal{A}x, x \rangle > 0 \quad \text{for all } 0 \neq x \in \mathbf{E}.$$

– A bilinear form $b(\cdot, \cdot)$ on the Euclidean space $(\mathbf{E}, \langle \cdot, \cdot \rangle)$ is an inner product on $\mathbf{E}$ if and only if there is a positive definite linear operator $\mathcal{A}$ on $\mathbf{E}$ such that $b(x, y) = \langle \mathcal{A}x, y \rangle \; \forall \, x, y \in \mathbf{E}$.

Positive Semi-Definite Linear Operators

A self-adjoint operator $\mathcal{A}$ is *positive semi-definite*, denoted $\mathcal{A} \succeq 0$, whenever

$$\langle \mathcal{A}x, x\rangle \geq 0 \quad \text{for all } x \in \mathbf{E}.$$

Similarly, a self-adjoint operator $\mathcal{A}$ is *positive definite*, denoted $\mathcal{A} \succ 0$, whenever

$$\langle \mathcal{A}x, x\rangle > 0 \quad \text{for all } 0 \neq x \in \mathbf{E}.$$

– A bilinear form $b(\cdot,\cdot)$ on the Euclidean space $(\mathbf{E}, \langle\cdot,\cdot\rangle)$ is an inner product on $\mathbf{E}$ if and only if there is a positive definite linear operator $\mathcal{A}$ on $\mathbf{E}$ such that $b(x, y) = \langle \mathcal{A}x, y\rangle \; \forall\, x, y \in \mathbf{E}$.

– For any two linear operators $\mathcal{A}$ and $\mathcal{B}$, we will use the notation $\mathcal{A} \succeq \mathcal{B}$ to mean $\mathcal{A} - \mathcal{B} \succeq 0$. The notation $\mathcal{A} \succ \mathcal{B}$ is defined similarly.

Norms

A *norm* on a vector space $\mathcal{V}$ is a function $\|\cdot\|: \mathcal{V} \to \mathbf{R}$ for which the following three properties hold for all point $x, y \in \mathcal{V}$ and scalars $a \in \mathbf{R}$:

(Absolute homogeneity) $\|ax\| = |a| \cdot \|x\|$

(Triangle inequality) $\|x + y\| \leq \|x\| + \|y\|$

(Positivity) Equality $\|x\| = 0$ holds if and only if $x = 0$.

The inner product in the Euclidean space $\mathbf{E}$ always induces a norm $\|x\| = \sqrt{\langle x, x \rangle}$. Unless specified otherwise, the symbol $\|x\|$ for $x \in \mathbf{E}$ will always denote this induced norm.

Examples of Norms

p-norms on $\mathbf{R}^n$:

$$\|x\|_p = \begin{cases} (|x_1|^p + \ldots + |x_n|^p)^{1/p} & \text{for } 1 \leq p < \infty \\ \max\{|x_1|, \ldots, |x_n|\} & \text{for } p = \infty \end{cases} .$$

Elliptic or inner product norms on $\mathbf{R}^n$: Let $\mathcal{A} \in \mathbf{L}(\mathbf{E}, \mathbf{E})$ be positive definite.

$$\|x\|_{\mathcal{A}} := \sqrt{\langle Ax, y\rangle}$$

Dual norms: Given an arbitrary norm $\|\cdot\|$ on $\mathbf{R}^n$, the norm dual to $\|\cdot\|$ is defined by

$$\|v\|^* := \max\{\langle v, x\rangle : \|x\| \leq 1\}.$$

Examples of Norms

p-norms on $\mathbf{R}^n$:

$$\|x\|_p = \begin{cases} (|x_1|^p + \ldots + |x_n|^p)^{1/p} & \text{for } 1 \le p < \infty \\ \max\{|x_1|, \ldots, |x_n|\} & \text{for } p = \infty \end{cases} .$$

Elliptic or inner product norms on $\mathbf{R}^n$: Let $\mathcal{A} \in \mathbf{L}(\mathbf{E}, \mathbf{E})$ be positive definite.

$$\|x\|_{\mathcal{A}} := \sqrt{\langle Ax, y \rangle}$$

Dual norms: Given an arbitrary norm $\|\cdot\|$ on $\mathbf{R}^n$, the norm dual to $\|\cdot\|$ is defined by

$$\|v\|^* := \max\{\langle v, x \rangle : \|x\| \le 1\}.$$

Why do norms and their duals satisfy the generalized Cauchy-Schwartz inequality

$$|\langle x, y \rangle| \le \|x\| \cdot \|y\|^* \qquad \text{for all } x, y \in \mathbf{E}\,?$$

Equivalence of Norms

All norms on $\mathbf{E}$ are "equivalent" in the sense that for any two norms $\rho_1(\cdot)$ and $\rho_2(\cdot)$, there exist constants $\alpha, \beta > 0$ satisfying

$$\alpha\rho_1(x) \le \rho_2(x) \le \beta\rho_1(x) \qquad \text{for all } x \in \mathbf{E}.$$

$$\begin{aligned} \|x\|_2 &\le \|x\|_1 \le \sqrt{n}\|x\|_2 \\ \|x\|_\infty &\le \|x\|_2 \le \sqrt{n}\|x\|_\infty \\ \|x\|_\infty &\le \|x\|_1 \le n\|x\|_\infty. \end{aligned}$$

The term "equivalent" is a misnomer since the constants α, β strongly depend on the (often enormous) dimension of the vector space $\mathbf{E}$. Hence measuring quantities in different norms can yield strikingly different conclusions.

The Orthogonal Group

Let $(\mathbf{E}, \langle\cdot,\cdot\rangle)$ be a Euclidean space. A linear operator $U \in \mathbf{L}(\mathbf{E},\mathbf{E})$ is said to be *distance preserving* if

$$\|Ux\| = \|x\| \;\; \forall\, x \in \mathbf{E},$$

where $\|x\| = \sqrt{\langle x.x\rangle}$ is the inner product norm on $\mathbf{E}$.

The set $\mathcal{O}(\mathbf{E})$ of all distance preserving linear operators on $\mathbf{E}$ is called the *orthogonal group* for $\mathbf{E}$, and the elements of $\mathcal{O}(\mathbf{E})$ are called *orthogonal operators*.

The Orthogonal Group

Let $(\mathbf{E}, \langle\cdot,\cdot\rangle)$ be a Euclidean space. A linear operator $U \in \mathbf{L}(\mathbf{E}, \mathbf{E})$ is said to be *distance preserving* if

$$\|Ux\| = \|x\| \ \forall x \in \mathbf{E},$$

where $\|x\| = \sqrt{\langle x.x\rangle}$ is the inner product norm on $\mathbf{E}$.

The set $\mathcal{O}(\mathbf{E})$ of all distance preserving linear operators on $\mathbf{E}$ is called the *orthogonal group* for $\mathbf{E}$, and the elements of $\mathcal{O}(\mathbf{E})$ are called *orthogonal operators*.

– $\mathcal{O}(\mathbf{E})$ is a group under matrix multiplication where the inverse of any element is simply its adjoint.

The Orthogonal Group

Let $(\mathbf{E}, \langle\cdot,\cdot\rangle)$ be a Euclidean space. A linear operator $U \in \mathbf{L}(\mathbf{E}, \mathbf{E})$ is said to be *distance preserving* if

$$\|Ux\| = \|x\| \;\; \forall\, x \in \mathbf{E},$$

where $\|x\| = \sqrt{\langle x.x\rangle}$ is the inner product norm on $\mathbf{E}$.
The set $\mathcal{O}(\mathbf{E})$ of all distance preserving linear operators on $\mathbf{E}$ is called the *orthogonal group* for $\mathbf{E}$, and the elements of $\mathcal{O}(\mathbf{E})$ are called *orthogonal operators*.

– $\mathcal{O}(\mathbf{E})$ is a group under matrix multiplication where the inverse of any element is simply its adjoint.

– Given a basis for $\mathbf{E}$, we can identify $\mathbf{L}(\mathbf{E}, \mathbf{E})$ with $\mathbf{R}^{n\times n}$ where n is the dimension of $\mathbf{E}$. If we identify $\mathcal{O}(\mathbf{E})$ with its associated matrices, then $\mathcal{O}(\mathbf{E}) = \{U \in \mathbf{R}^{n\times n} \,|\, UU^T = I = U^TU\}$ and its elements are called *orthogonal matrices*.

Eigenvalues of Symmetric Matrices

Let $A \in \mathbf{S}^n$. $\lambda \in \mathbf{R}$ is an eigenvalue for A if $exists\, x \in \mathbf{R}^n \setminus \{0\}$ s.t. $Ax = \lambda x$.

The vector x is called an eigenvector associated with λ.

Eigenvalues of Symmetric Matrices

Let $A \in \mathbf{S}^n$. $\lambda \in \mathbf{R}$ is an eigenvalue for A if *exists* $x \in \mathbf{R}^n \setminus \{0\}$ s.t. $Ax = \lambda x$.

The vector x is called an eigenvector associated with λ.

Note $x \in \ker(A - \lambda I)$, where,

$\forall\, B \in \mathbf{R}^{n\times n},\ \ker B := \{w \in \mathbf{R}^n | Bw = 0\}$.

Eigenvalues of Symmetric Matrices

Let $A \in \mathbf{S}^n$. $\lambda \in \mathbf{R}$ is an eigenvalue for A if *exists* $x \in \mathbf{R}^n \setminus \{0\}$ s.t. $Ax = \lambda x$.

The vector x is called an eigenvector associated with λ.

Note $x \in \ker(A - \lambda I)$, where,

$\forall\, B \in \mathbf{R}^{n\times n},\ \ker B := \{w \in \mathbf{R}^n | Bw = 0\}$.

Consequently, the eigenvalues of A are the roots of the *characteristic polynomial*

$$\lambda \mapsto \det(A - \lambda I).$$

Eigenvalues of Symmetric Matrices

Let $A \in \mathbf{S}^n$. $\lambda \in \mathbf{R}$ is an eigenvalue for A if *exists* $x \in \mathbf{R}^n \setminus \{0\}$ s.t. $Ax = \lambda x$.

The vector x is called an eigenvector associated with λ.

Note $x \in \ker(A - \lambda I)$, where,

$\forall\, B \in \mathbf{R}^{n\times n}$, $\ker B := \{w \in \mathbf{R}^n | Bw = 0\}$.

Consequently, the eigenvalues of A are the roots of the *characteristic polynomial*

$$\lambda \mapsto \det(A - \lambda I).$$

If $A \in \mathbf{S}^n$, these n roots are real. One can show that there is an associated orthonormal basis of real eigenvectors. Consequently, A is diagonalizable in the sense that

$$U^T A U = \Lambda \qquad \text{or} \qquad A = U \Lambda U^T,$$

where the columns of $U \in \mathcal{O}(\mathbf{R}^n)$ are an orthonormal basis of eigenvectors and Λ is the diagonal matrix of corresponding eigenvalues.

Rayleigh-Ritz Theorem and Square Roots

Fix an ordering and denote the eigenvalues of A by

$$\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A).$$

A simple consequence of the decomposition $A = U\Lambda U^T$ is the Rayleigh-Ritz theorem:

$$\lambda_n(A) \leq \frac{\langle Au, u\rangle}{\langle u, u\rangle} \leq \lambda_1(A) \quad \text{for all } u \in \mathbf{R}^n \setminus \{0\}.$$

Rayleigh-Ritz Theorem and Square Roots

Fix an ordering and denote the eigenvalues of A by

$$\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A).$$

A simple consequence of the decomposition $A = U\Lambda U^T$ is the Rayleigh-Ritz theorem:

$$\lambda_n(A) \leq \frac{\langle Au, u\rangle}{\langle u, u\rangle} \leq \lambda_1(A) \quad \text{for all } u \in \mathbf{R}^n \setminus \{0\}.$$

Observe that the two conditions, $A \succeq 0$ and $\lambda_n(A) \geq 0$ are equivalent; similarly, $A \succ 0$ if and only $\lambda_n(A) > 0$.

Rayleigh-Ritz Theorem and Square Roots

Fix an ordering and denote the eigenvalues of A by

$$\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A).$$

A simple consequence of the decomposition $A = U\Lambda U^T$ is the Rayleigh-Ritz theorem:

$$\lambda_n(A) \leq \frac{\langle Au, u\rangle}{\langle u, u\rangle} \leq \lambda_1(A) \quad \text{for all } u \in \mathbf{R}^n \setminus \{0\}.$$

Observe that the two conditions, $A \succeq 0$ and $\lambda_n(A) \geq 0$ are equivalent; similarly, $A \succ 0$ if and only $\lambda_n(A) > 0$.

Consequently, $A \in \mathbf{S}^n$ is positive semidefinite if and only if there exists a matrix $B \in \mathbf{S}^n$ satisfying $A = BB^T$ (why?). The matrix B is called a *square root* of A. There are infinitely many such such square roots (see *Cholesky Factorizations*). The spectral square root is $B = U\Lambda^{1/2}U^T =: \sqrt{A}$.

The Singular Value Decomposition

Given $A, B^T \in \mathbf{R}^{m\times n}$, one can show that the nonzero eigenvalues of AB coincide with those of BA including multiplicity.

Therefore, the eigenvalues of the symmetric matrices A^TA and AA^T coincide up to multiplicity. Since these matrices are positive semi-definite (why?), their nonzero eigenvalues are positives and coincide up to multiplicity.

Let $k := \min\{n, m\}$ and define

$$\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq \sigma_k(A) \geq 0$$

to be the largest k eigenvalues of $\sqrt{A^TA}$ and note that any other eigenvalue of $\sqrt{A^TA}$ must be zero. The σ_is are called the *singular values of* A.

The Singular Value Decomposition

If the columns of $V \in \mathcal{O}(\mathbf{R}^n)$ form an orthonormal basis of eigenvectors for A^TA ordered in correspondence with the magnitude of its eigenvalues, it can be shown that there is a corresponding $U \in \mathcal{O}(\mathbf{R}^m)$ whose columns form an orthonormal basis of eigenvectors for AA^T such that

$$A = U\Sigma V^T,$$

where the principal diagonal $\Sigma \in \mathbf{R}^{m\times n}$ are the ordered singular values of A with all other values zero.

The Singular Value Decomposition

If the columns of $V \in \mathcal{O}(\mathbf{R}^n)$ form an orthonormal basis of eigenvectors for $A^T A$ ordered in correspondence with the magnitude of its eigenvalues, it can be shown that there is a corresponding $U \in \mathcal{O}(\mathbf{R}^m)$ whose columns form an orthonormal basis of eigenvectors for AA^T such that

$$A = U\Sigma V^T,$$

where the principal diagonal $\Sigma \in \mathbf{R}^{m\times n}$ are the ordered singular values of A with all other values zero.

If we let $k := \text{rank}(A)$, we may write

$$A = U\Sigma V^T,$$

where now $U \in \mathbf{R}^{m\times k},\ V \in \mathbf{R}^{n\times k}$ have orthogonal columns and $\Sigma \in \mathbf{R}^{k\times k}$ is diagonal with the ordered nonzero singular values on the diagonal. This called the compact or reduced singular value decomposition.

The Singular Value Decomposition

If the columns of $V \in \mathcal{O}(\mathbf{R}^n)$ form an orthonormal basis of eigenvectors for $A^T A$ ordered in correspondence with the magnitude of its eigenvalues, it can be shown that there is a corresponding $U \in \mathcal{O}(\mathbf{R}^m)$ whose columns form an orthonormal basis of eigenvectors for AA^T such that

$$A = U \Sigma V^T,$$

where the principal diagonal $\Sigma \in \mathbf{R}^{m \times n}$ are the ordered singular values of A with all other values zero.

If we let $k := \operatorname{rank}(A)$, we may write

$$A = U \Sigma V^T,$$

where now $U \in \mathbf{R}^{m \times k}$, $V \in \mathbf{R}^{n \times k}$ have orthogonal columns and $\Sigma \in \mathbf{R}^{k \times k}$ is diagonal with the ordered nonzero singular values on the diagonal. This called the compact or reduced singular value decomposition.

For this reason, some authors refer to only the nonzero singular values as the singular values. The columns of U are called the *left singular vectors* and those of V are the *right singular vectors*.

The Operator Norm on $\mathbf{R}^{n\times n}$

The Rayleigh-Ritz Theorem tells us that

$$\|A\|_{\mathrm{op}} := \sup_{x:\|x\|\leq 1} \|Ax\| \;=\; \sigma_1(A),$$

where $\|A\|_{\mathrm{op}}$ is called the *operator norm* of A when the given norms are the inner product norms.

The Operator Norm on $\mathbf{R}^{n\times n}$

The Rayleigh-Ritz Theorem tells us that

$$\|A\|_{\text{op}} := \sup_{x:\|x\|\le 1} \|Ax\| \; = \; \sigma_1(A),$$

where $\|A\|_{\text{op}}$ is called the *operator norm* of A when the given norms are the inner product norms.

Let $\sigma : \mathbf{R}^{m\times n} \to \mathbf{R}^k$, where $k := \min\{m, n\}$, be the mapping that takes a matrix to its ordered vector of singular values:

$$\sigma(A) := (\sigma_1(A), \sigma_2(A), \ldots, \sigma_k(A))^T.$$

The *Schatten p-norm* of a $A \in \mathbf{R}^{m\times n}$, for $1 \le p \le \infty$ is given by

$$\|A\|_p := \|\sigma(A)\|_p \, .$$

Hence $\|A\|_{\text{op}} = \|\sigma(A)\|_\infty$. For $p = 1$, $\|A\|_1$ is called the *nuclear* or *trace* norm.

The Operator Norm on $\mathbf{R}^{n\times n}$

The Rayleigh-Ritz Theorem tells us that

$$\|A\|_{\text{op}} := \sup_{x:\|x\|\le 1} \|Ax\| \;=\; \sigma_1(A),$$

where $\|A\|_{\text{op}}$ is called the *operator norm* of A when the given norms are the inner product norms.

Let $\sigma : \mathbf{R}^{m\times n} \to \mathbf{R}^k$, where $k := \min\{m, n\}$, be the mapping that takes a matrix to its ordered vector of singular values:

$$\sigma(A) := (\sigma_1(A), \sigma_2(A), \ldots, \sigma_k(A))^T.$$

The *Schatten p-norm* of a $A \in \mathbf{R}^{m\times n}$, for $1 \le p \le \infty$ is given by

$$\|A\|_p := \|\sigma(A)\|_p \,.$$

Hence $\|A\|_{\text{op}} = \|\sigma(A)\|_\infty$. For $p = 1$, $\|A\|_1$ is called the *nuclear* or *trace* norm.

It can be shown that all of the Schatten p-norms are norms on the Euclidean space $\mathbf{R}^{m\times n}$.

Sets and Operations on Sets

Let $\mathbf{X}$, $\mathbf{Y}$ be Euclidean spaces with $X_i \subset \mathbf{X}$ $i = 1, 2$, $Y \subset \mathbf{Y}$, and let $\mathcal{A} \in \mathbf{L}(\mathbf{X}, \mathbf{Y})$.

- $\mathbf{R}^n_+ := \{x \in \mathbf{R}^n | x_i \geq 0,\ i = 1, \ldots, n\}$, $\mathbf{R}^n_{++} := \{x \in \mathbf{R}^n | x_i > 0,\ i = 1, \ldots, n\}$
- $\mathbf{S}^n{}_+ := \{H \in \mathbf{S}^n \,|\, H \succeq 0\}$, $\mathbf{S}^n{}_{++} := \{H \in \mathbf{S}^n \,|\, H \succ 0\}$
- For $\lambda \in \mathbf{R}$, $\lambda X := \{\lambda x \,|\, x \in X\}$.
- $X_1 + X_2 := \left\{x^1 + x^2 \,|\, x_i \in X_i,\ i = 1, 2\right\}$ with $X_1 - X_2$ defined similarly.
- $\mathbf{R}_+ Y := \{\lambda y \,|\, \lambda \in \mathbf{R}_+,\ y \in Y\}$, the cone generated by Y.
- An *affine* set is a translate of a subspace.
- The affine hull of Y, $\operatorname{aff} Y$, is the intersection of all affine sets containing Y.
- $\mathcal{A} X_1 := \{\mathcal{A} x \,|\, x \in X_1\}$
- $\mathcal{A}^{-1} Y := \{x \,|\, \mathcal{A} x \in Y\}$
- $X_1 \times Y := \{(x, y) \,|\, x \in X_1,\ y \in Y\}$

Convex Sets

A set $C \subset \mathbf{E}$ is said to be convex if

$$x, y \in C \text{ and } \lambda \in [0,1] \implies (1-\lambda)x + \lambda y \in C.$$

That is, C contains all line segments connecting points in C.

Convex Sets

A set $C \subset \mathbf{E}$ is said to be convex if

$$x, y \in C \text{ and } \lambda \in [0,1] \implies (1-\lambda)x + \lambda y \in C.$$

That is, C contains all line segments connecting points in C.

Let $\mathcal{A} \in \mathbf{L}(\mathbf{E}, \mathbf{Y})$. If $C_1, C_2 \subset \mathbf{E}$ and $K \subset \mathbf{Y}$ are all convex, then so are the sets

– $C_1 \cap C_2$

– $\mathbf{R}_+ K$ and $\lambda K \quad \forall \lambda \in \mathbf{R}$

– $C_1 + C_2$

– $\mathcal{A}C_1$ and $\mathcal{A}^{-1}K$

– $C_1 \times K$

– $\mathrm{cl}\, K$ and $\mathrm{intr} K$

Convex Sets

A set $C \subset \mathbf{E}$ is said to be convex if

$$x, y \in C \text{ and } \lambda \in [0,1] \implies (1-\lambda)x + \lambda y \in C.$$

That is, C contains all line segments connecting points in C.

Let $\mathcal{A} \in \mathbf{L}(\mathbf{E}, \mathbf{Y})$. If $C_1, C_2 \subset \mathbf{E}$ and $K \subset \mathbf{Y}$ are all convex, then so are the sets

- $C_1 \cap C_2$
- $\mathbf{R}_+ K$ and $\lambda K \quad \forall\, \lambda \in \mathbf{R}$
- $C_1 + C_2$
- $\mathcal{A} C_1$ and $\mathcal{A}^{-1} K$
- $C_1 \times K$
- $\operatorname{cl} K$ and $\operatorname{intr} K$

We will spend a lot of time with convex sets.

Point-Set Topology

Let $\mathbf{E}$ be a Euclidean space with $x \in X \subset \mathbf{E}$.

- Given $r > 0$, the open r ball around x is the set
$$B_r(x) := \{y \mid \|x - y\| < r\}.$$
- x is in the closure of X, written $x \in \operatorname{cl} X$, if $B_r(x) \cap X \neq \emptyset \ \forall\, r > 0$.
- X is closed if $X = \operatorname{cl} X$.
- $x \in X$ is in the interior of X, written $x \in \operatorname{intr} X$, if there is an $r > 0$ such that $B_r(x) \subset X$.
- X is open if $X = \operatorname{intr} X$.
- X is bounded if there is an $r > 0$ such that $X \subset B_r(0)$.
- X is compact if it is closed and bounded.

Point-Set Topology

Let $\mathbf{E}$ be a Euclidean space with $x \in X \subset \mathbf{E}$.

- Given $r > 0$, the open r ball around x is the set
$$B_r(x) := \{y \mid \|x - y\| < r\}.$$
- x is in the closure of X, written $x \in \operatorname{cl} X$, if $B_r(x) \cap X \neq \emptyset \ \forall \, r > 0$.
- X is closed if $X = \operatorname{cl} X$.
- $x \in X$ is in the interior of X, written $x \in \operatorname{intr} X$, if there is an $r > 0$ such that $B_r(x) \subset X$.
- X is open if $X = \operatorname{intr} X$.
- X is bounded if there is an $r > 0$ such that $X \subset B_r(0)$.
- X is compact if it is closed and bounded.

Theorem (Bolzano-Weierstrass)

$Q \subset \mathbf{E}$ is compact if and only if every sequence in Q admits a subsequence converging to a point in Q.

Limits Inferior and Superior

Define the *extended real line* $\overline{\mathbf{R}} := \mathbf{R} \cup \{\pm\infty\}$.

The *limit inferior* and *limit superior* of any sequence $\{r_i\} \subset \overline{\mathbf{R}}$ are defined by

$$\liminf_{i\to\infty} r_i = \lim_{i\to\infty} \left\{ \inf_{j\geq i} r_j \right\} \qquad \text{and} \qquad \limsup_{i\to\infty} r_i = \lim_{i\to\infty} \left\{ \sup_{j\geq i} r_j \right\}.$$

For any function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ and a point $x \in \mathbf{E}$, we set

$$\liminf_{y\to x} f(y) = \lim_{r>0} \left\{ \inf_{y\in B_r(x)\setminus\{x\}} f(y) \right\}$$

The symbol $\limsup_{y\to x} f(y)$ is defined similarly, with sup replacing inf.

Limits Inferior and Superior

Define the *extended real line* $\overline{\mathbf{R}} := \mathbf{R} \cup \{\pm\infty\}$.

The *limit inferior* and *limit superior* of any sequence $\{r_i\} \subset \overline{\mathbf{R}}$ are defined by

$$\liminf_{i\to\infty} r_i = \lim_{i\to\infty} \left\{ \inf_{j\geq i} r_j \right\} \qquad \text{and} \qquad \limsup_{i\to\infty} r_i = \lim_{i\to\infty} \left\{ \sup_{j\geq i} r_j \right\}.$$

For any function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ and a point $x \in \mathbf{E}$, we set

$$\liminf_{y\to x} f(y) = \lim_{r>0} \left\{ \inf_{y\in B_r(x)\setminus\{x\}} f(y) \right\}$$

The symbol $\limsup_{y\to x} f(y)$ is defined similarly, with sup replacing inf.

Note: The infimum (supremum) over the empty set is $+\infty$ $(-\infty)$.

Functions and Continuity

Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ and $F : \mathbf{X} \to \mathbf{Y}$.

- $\operatorname{dom} f := \{x \mid f(x) < \infty\}$
- $\operatorname{epi} f := \{(x, \lambda) \mid f(x) \le \lambda\} \subset \mathbf{E} \times \mathbf{R}$
- f is lower semi-continuous (lsc) at $x \in \mathbf{E}$ if $\liminf_{y \to x} f(y) \ge f(x)$. f is *closed* if it is lsc for all $x \in \mathbf{E}$.
- f is upper semi-continuous at $x \in \mathbf{E}$ if $f(x) \ge \limsup_{y \to x} f(y)$.
- f is continuous at $x \in \operatorname{intr}(\operatorname{dom} f)$ if $\liminf_{y \to x} f(y) = f(x) = \limsup_{y \to x} f(y)$.
- F continuous at $x \in \mathbf{X}$ if
$$\forall \epsilon > 0 \ \exists \delta > 0 \text{ s.t. } \|F(y) - F(x)\| \le \epsilon \text{ when } \|y - x\| \le \delta.$$
- For $L > 0$, F is L-Lipschitz continuous at $x \in \mathbf{X}$ if
$$\|F(x) - F(y)\| \le L \|x - y\|.$$
- For $L > 0$ and $X \subset \mathbf{X}$, F is L-Lipschitz continuous on X if it is L-Lipschitz continuous forall $x \in \mathbf{X}$. If $X = \mathbf{X}$, we simply say F is L-Lipschitz. If $0 < L < 1$, we say that F is a *contraction*.

Functions and Continuity

Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ and $F : \mathbf{X} \to \mathbf{Y}$.

- $\operatorname{dom} f := \{x \mid f(x) < \infty\}$
- $\operatorname{epi} f := \{(x, \lambda) \mid f(x) \leq \lambda\} \subset \mathbf{E} \times \mathbf{R}$
- f is lower semi-continuous (lsc) at $x \in \mathbf{E}$ if $\liminf_{y \to x} f(y) \geq f(x)$. f is *closed* if it is lsc for all $x \in \mathbf{E}$.
- f is upper semi-continuous at $x \in \mathbf{E}$ if $f(x) \geq \limsup_{y \to x} f(y)$.
- f is continuous at $x \in \operatorname{intr}(\operatorname{dom} f)$ if $\liminf_{y \to x} f(y) = f(x) = \limsup_{y \to x} f(y)$.
- F continuous at $x \in \mathbf{X}$ if
$$\forall \epsilon > 0 \; \exists \delta > 0 \text{ s.t. } \|F(y) - F(x)\| \leq \epsilon \text{ when } \|y - x\| \leq \delta.$$
- For $L > 0$, F is L-Lipschitz continuous at $x \in \mathbf{X}$ if
$$\|F(x) - F(y)\| \leq L \|x - y\|.$$
- For $L > 0$ and $X \subset \mathbf{X}$, F is L-Lipschitz continuous on X if it is L-Lipschitz continuous forall $x \in \mathbf{X}$. If $X = \mathbf{X}$, we simply say F is L-Lipschitz. If $0 < L < 1$, we say that F is a *contraction*.

Theorem

$f : \mathbf{E} \to \overline{\mathbf{R}}$ is closed if and only if $\operatorname{epi} f$ is closed.

Existence of Optimal Solutions

Theorem (Weierstrass Extrema Value Theorem)

A continuous function on a compact set attains its extrema values on that set. That is, if $f : C \to \mathbf{R}$ is continuous on the compact set $C \subset \mathbf{E}$, then there exist $\bar{x}, \overline{y} \in C$ such that $f(\bar{x}) \leq f(x) \leq f(\overline{y})$ for all $x \in C$.

This can be refined using lower semi-continuity.

Theorem

If $f : Q \to \mathbf{R}$ is closed with $Q \subset \mathbf{E}$ compact, then there is an $\bar{x} \in Q$ such that $f(\bar{x}) \leq f(x)$ for all $x \in Q$.

Existence of Optimal Solutions

Theorem (Weierstrass Extrema Value Theorem)

A continuous function on a compact set attains its extrema values on that set. That is, if $f : C \to \mathbf{R}$ is continuous on the compact set $C \subset \mathbf{E}$, then there exist $\bar{x}, \bar{y} \in C$ such that $f(\bar{x}) \le f(x) \le f(\bar{y})$ for all $x \in C$.

This can be refined using lower semi-continuity.

Theorem

If $f : Q \to \mathbf{R}$ is closed with $Q \subset \mathbf{E}$ compact, then there is an $\bar{x} \in Q$ such that $f(\bar{x}) \le f(x)$ for all $x \in Q$.

Coercive Functions: A function $f : \mathbf{E} \to \overline{\mathbf{R}}$ is *coercive* if for any sequence x_i with $\|x_i\| \to \infty$, it must be that $f(x_i) \to +\infty$.

It is easy to show that f is coercive if and only if the sets $\{x \,|\, f(x) \le r\}$ are compact for all $r \in \mathbf{R}$. This observation implies that any closed coercive function has a global minimizer, i.e. there is $\bar{x}$ such that $f(\bar{x}) \le f(x)$ for all $x \in \mathbf{E}$.

Linear Operators

Let $\mathbf{X}$ and $\mathbf{Y}$ be real normed linear spaces with norms $\|\cdot\|_x$ and $\|\cdot\|_y$, respectively.
A linear transformation (or operator) from $\mathbf{X}$ to $\mathbf{Y}$ is any mapping $\mathcal{L}: \mathbf{X} \to \mathbf{Y}$ such that

$$\mathcal{L}(\alpha x + \beta z) = \alpha\mathcal{L}(x) + \beta\mathcal{L}(z) \quad \forall\, x, z \in \mathbf{X} \text{ and } \alpha, \beta \in \mathbf{R}.$$

The linear operator $\mathcal{T}$ is continuous with respect to the norms on $\mathbf{X}$ and $\mathbf{Y}$ if and only if

$$\|\mathcal{T}\| := \sup_{\|x\|_x \leq 1} \|\mathcal{T}x\|_y \quad \forall\, \mathcal{T} \in \mathbf{L}[\mathbf{X}, \mathbf{Y}],$$

is finite.
Let $\mathbf{L}[\mathbf{X}, \mathbf{Y}]$ denote the space of all continuous linear operators from $\mathbf{X}$ to $\mathbf{Y}$. In can be shown that $\|\mathcal{T}\|$ is a norm on this space.

Linear Operators

Let $\mathbf{X}$ and $\mathbf{Y}$ be real normed linear spaces with norms $\|\cdot\|_x$ and $\|\cdot\|_y$, respectively.
A linear transformation (or operator) from $\mathbf{X}$ to $\mathbf{Y}$ is any mapping $\mathcal{L} : \mathbf{X} \to \mathbf{Y}$ such that

$$\mathcal{L}(\alpha x + \beta z) = \alpha\mathcal{L}(x) + \beta\mathcal{L}(z) \quad \forall\, x, z \in \mathbf{X} \text{ and } \alpha, \beta \in \mathbf{R}.$$

The linear operator $\mathcal{T}$ is continuous with respect to the norms on $\mathbf{X}$ and $\mathbf{Y}$ if and only if

$$\|\mathcal{T}\| := \sup_{\|x\|_x \leq 1} \|\mathcal{T}x\|_y \quad \forall\, \mathcal{T} \in \mathbf{L}[\mathbf{X}, \mathbf{Y}],$$

is finite.
Let $\mathbf{L}[\mathbf{X}, \mathbf{Y}]$ denote the space of all continuous linear operators from $\mathbf{X}$ to $\mathbf{Y}$. In can be shown that $\|\mathcal{T}\|$ is a norm on this space.
The topological dual of the normed linear space $\mathbf{X}$ is

$$\mathbf{X}^* := \mathbf{L}[\mathbf{X}, \mathbf{R}]$$

with the *duality pairing* denoted by

$$\langle \phi,\, x\rangle = \phi(x) \quad \forall\, (\phi, x) \in \mathbf{X}^* \times \mathbf{X}.$$

Hilbert Spaces

If the norm on $\mathbf{X}$ satisfies the parallelogram law,

$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2\,,$$

then we call $\mathbf{X}$ a Hilbert space.

In this case there of a natural isometry between $\mathbf{X}^*$ and $\mathbf{X}$ under which the duality pairing is an inner product:

$$\langle x, y\rangle = \frac{\|x+y\|^2 - \|x-y\|^2}{4}.$$

Hilbert Spaces

If the norm on $\mathbf{X}$ satisfies the parallelogram law,

$$\|x - y\|^2 + \|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2 ,$$

then we call $\mathbf{X}$ a Hilbert space.

In this case there of a natural isometry between $\mathbf{X}^*$ and $\mathbf{X}$ under which the duality pairing is an inner product:

$$\langle x, y\rangle = \frac{\|x + y\|^2 - \|x - y\|^2}{4}.$$

Note: A Euclidean space is a real finite dimensional Hilbert space.

Bilinear Forms

Let $\mathbf{X}$ and $\mathbf{Y}$ be real linear spaces. A mapping $\mathcal{Q} : \mathbf{X} \times \mathbf{X} \to \mathbf{Y}$ is said to be a bilinear if it is linear in each argument separately: for all $(x^i, z^j) \in \mathbf{X} \times \mathbf{X}$, $i = 1, 2$, and $\alpha, \beta, \gamma, \delta \in \mathbf{R}$

$$\begin{aligned}\mathcal{Q}(\alpha x^1 + \beta x^2, \gamma z^1 + \delta z^2) &= \alpha \mathcal{Q}(x^1, \gamma z^1 + \delta z^2) + \beta \mathcal{Q}(x^2, \gamma z^1 + \delta z^2) \\ &= \gamma \mathcal{Q}(\alpha x^1 + \beta x^2, z^1) + \delta \mathcal{Q}(\alpha x^1 + \beta x^2, z^2).\end{aligned}$$

The bilinear form $\mathcal{Q}$ is said to be symmetric if $\mathcal{Q}(x, z) = \mathcal{Q}(z, x)$.

Let $\mathbf{B}[\mathbf{X}, \mathbf{Y}]$ denote the set of all continuous bilinear maps from $\mathbf{X}$ to $\mathbf{Y}$.

If $\mathbf{Y} = \mathbf{R}$, the bilinear map $\mathcal{Q}$ is call a *bilinear form* and we write $\mathbf{Q}[\mathbf{X}] := \mathbf{B}[\mathbf{X}, \mathbf{R}]$.

Differentiability

Let $U \subset \mathbf{E}$ be open.
$f : U \to \mathbf{R}$ is *differentiable* at $x \in U$ if there exists a vector, denoted by $\nabla f(x) \in \mathbf{E}$, satisfying

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle \nabla f(x), h \rangle}{\|h\|} = 0.$$

We call $\nabla f(x)$ the *gradient* of f at x.
If $\mathbf{E} = \mathbf{R}^n$,

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}.$$

Differentiability

Let $U \subset \mathbf{E}$ be open.
$f : U \to \mathbf{R}$ is *differentiable* at $x \in U$ if there exists a vector, denoted by $\nabla f(x) \in \mathbf{E}$, satisfying

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle \nabla f(x), h \rangle}{\|h\|} = 0.$$

We call $\nabla f(x)$ the *gradient* of f at x.
If $\mathbf{E} = \mathbf{R}^n$,

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}.$$

Let the symbol $o(r)$ represent the class of functions satisfying $0 = \lim_{r \downarrow 0} o(r)/r$. Then f is differentiable at x if and only if

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + o(\|h\|).$$

Differentiability

If the mapping $\nabla f : U \to \mathbf{R}^n$ is well-defined and continuous, we say f is $\mathcal{C}^1$-*smooth* on U.

If the gradient satisfies the stronger Lipschitz property

$$\|\nabla f(y) - \nabla f(x)\| \leq \beta \|y - x\| \qquad \text{holds for all } x, y \in U,$$

then we say that f is β-*smooth*.

Differentiability

If the mapping $\nabla f : U \to \mathbf{R}^n$ is well-defined and continuous, we say f is $\mathcal{C}^1$*-smooth* on U.

If the gradient satisfies the stronger Lipschitz property

$$\|\nabla f(y) - \nabla f(x)\| \leq \beta \|y - x\| \qquad \text{holds for all } x, y \in U,$$

then we say that f is β*-smooth*.

More generally, a mapping $F \colon U \to \mathbf{Y}$ is *differentiable* at $x \in U$ if there exists a linear mapping from $\mathbf{E}$ to $\mathbf{Y}$, denoted by $F'(x)$, satisfying

$$F(x + h) = F(x) + F'(x)h + o(\|h\|).$$

Differentiability

If the mapping $\nabla f : U \to \mathbf{R}^n$ is well-defined and continuous, we say f is $\mathcal{C}^1$-*smooth* on U.

If the gradient satisfies the stronger Lipschitz property

$$\|\nabla f(y) - \nabla f(x)\| \leq \beta \|y - x\| \qquad \text{holds for all } x, y \in U,$$

then we say that f is β-*smooth*.

More generally, a mapping $F \colon U \to \mathbf{Y}$ is *differentiable* at $x \in U$ if there exists a linear mapping from $\mathbf{E}$ to $\mathbf{Y}$, denoted by $F'(x)$, satisfying

$$F(x + h) = F(x) + F'(x)h + o(\|h\|).$$

If one chooses bases in $\mathbf{E}$ and $\mathbf{Y}$, then $F'(x) \in \mathbf{L}(\mathbf{E}, \mathbf{Y})$ can be given a matrix representation which is denoted by $\nabla F(x)$ and called the *Jacobian* of F at x. If the assignment $x \mapsto F'(x)$ is continuous, we say that F is $\mathcal{C}^1$-*smooth*.

Differentiability

If $\mathbf{E} = \mathbf{R}^n$ and $\mathbf{Y} = \mathbf{R}^m$, we can write F in terms of coordinate functions $F(x) = (F_1(x), \ldots, F_m(x))$, and then the Jacobian is simply

$$\nabla F(x) = \begin{pmatrix} \nabla F_1(x)^T \\ \nabla F_2(x)^T \\ \vdots \\ \nabla F_m(x)^T \end{pmatrix} = \begin{pmatrix} \frac{\partial F_1(x)}{\partial x_1} & \frac{\partial F_1(x)}{\partial x_2} & \cdots & \frac{\partial F_1(x)}{\partial x_n} \\ \frac{\partial F_2(x)}{\partial x_1} & \frac{\partial F_2(x)}{\partial x_2} & \cdots & \frac{\partial F_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m(x)}{\partial x_1} & \frac{\partial F_m(x)}{\partial x_2} & \cdots & \frac{\partial F_m(x)}{\partial x_n} \end{pmatrix}.$$

Calculus Rules

Let $U \subset \mathbf{E}$ and $W \subset \mathbf{Y}$ be open.
Let $F_i : U \to \mathbf{Y}$, $i = 1, 2$, $F : U \to W$, and $H : W \to \mathbf{Z}$ be $\mathcal{C}^1$ (this can be significantly weakened).

– If $F \in \mathbf{L}(\mathbf{E}, \mathbf{Y})$, the $F'(x) = F$ for al $x \in \mathbf{E}$.

– For all $\lambda \in \mathbf{R}$ and $x \in U$, $'(\lambda F)'(x) = \lambda F'(x)$.

– For all $x \in U$, $(F_1 + F_2)'(x) = F_1'(x) + F_2'(x)$.

– **The Chain Rule:** The mapping $G : U \to \mathbf{Z}$ given by $G := H \circ F$ is differentiable on U with $G'(x) = H'(F(x)) \circ F'(x)$.

Example

Let $A \in \mathbf{R}^{s\times n}$ and $B \in \mathbf{R}^{n\times t}$ and consider the mapping $\mathcal{T} : \mathbf{R}^{m\times n} \to \mathbf{R}^{s\times t}$ given by

$$\mathcal{T}(X) := AXB.$$

Clearly, $\mathcal{T} \in \mathbf{L}(\mathbf{R}^{m\times n}, \mathbf{R}^{s\times t})$, hence

$$\mathcal{T}'(X)Y = \mathcal{T}(Y) = AYB \qquad \forall\ X \in \mathbf{R}^{m\times n}.$$

What is $\nabla\mathcal{T}$?

Example

Let $A \in \mathbf{R}^{s\times n}$ and $B \in \mathbf{R}^{n\times t}$ and consider the mapping $\mathcal{T} : \mathbf{R}^{m\times n} \to \mathbf{R}^{s\times t}$ given by

$$\mathcal{T}(X) := AXB.$$

Clearly, $\mathcal{T} \in \mathbf{L}(\mathbf{R}^{m\times n}, \mathbf{R}^{s\times t})$, hence

$$\mathcal{T}'(X)Y = \mathcal{T}(Y) = AYB \qquad \forall\ X \in \mathbf{R}^{m\times n}.$$

What is $\nabla\mathcal{T}$?

Representing the matrix $\nabla\mathcal{T}$ requires choosing bases in both $\mathbf{R}^{m\times n}$ and $\mathbf{R}^{s\times t}$ and then recording the action of $\mathcal{T}$ on these bases. This is doable, but it is a real mess. A helpful tool in this regard is the *Kronecker product* to be discussed later.

The Second Derivative

Let $F : \mathbf{X} \to \mathbf{Y}$ we say that F is twice differentiable at x if F is differentiable at x and there is a bilinear form $\mathcal{Q}$ such that

$$\lim_{z \to x} \frac{\left\| F(z) - (F(z) + \nabla F(x)(z-x) + \frac{1}{2}\mathcal{Q}(z-x, z-x)) \right\|}{\|y - x\|^2} = 0.$$

We call $\mathcal{Q}$ the second derivative of F at x and write $\mathcal{Q} = F''(x)$. If the mapping $x \to F''(x)$ is continuous, we say that F is $\mathcal{C}^2$.

The Second Derivative

Let $F : \mathbf{X} \to \mathbf{Y}$ we say that F is twice differentiable at x if F is differentiable at x and there is a bilinear form $\mathcal{Q}$ such that

$$\lim_{z \to x} \frac{\left\| F(z) - (F(z) + \nabla F(x)(z-x) + \frac{1}{2}\mathcal{Q}(z-x, z-x)) \right\|}{\left\| y - x \right\|^2} = 0.$$

We call $\mathcal{Q}$ the second derivative of F at x and write $\mathcal{Q} = F''(x)$. If the mapping $x \to F''(x)$ is continuous, we say that F is $\mathcal{C}^2$.

When $\mathbf{X} = \mathbf{R}^n$ and $\mathbf{Y} = \mathbf{R}$, we call $F''(x)$ the Hessian of F at x and write $\nabla^2 F(x) := F''(x)$. If all of the second partials of F are continuous, then $\nabla^2 F(x) \in \mathbf{S}^n$ is the $n \times n$ matrix of second partials.

The Second Derivative

Let $F : \mathbf{X} \to \mathbf{Y}$ we say that F is twice differentiable at x if F is differentiable at x and there is a bilinear form $\mathcal{Q}$ such that

$$\lim_{z \to x} \frac{\left\| F(z) - (F(z) + \nabla F(x)(z-x) + \frac{1}{2}\mathcal{Q}(z-x, z-x)) \right\|}{\|y-x\|^2} = 0.$$

We call $\mathcal{Q}$ the second derivative of F at x and write $\mathcal{Q} = F''(x)$. If the mapping $x \to F''(x)$ is continuous, we say that F is $\mathcal{C}^2$.

When $\mathbf{X} = \mathbf{R}^n$ and $\mathbf{Y} = \mathbf{R}$, we call $F''(x)$ the Hessian of F at x and write $\nabla^2 F(x) := F''(x)$. If all of the second partials of F are continuous, then $\nabla^2 F(x) \in \mathbf{S}^n$ is the $n \times n$ matrix of second partials.

Again, the *little-o* notation gives
$F(y) = F(x) + \langle \nabla F(x), (y-x) \rangle + \frac{1}{2} \langle \nabla^2 F(x)(y-x), (y-x) \rangle + o(\|y-x\|^2)$.

Computing Derivatives

Consider the linear transformation $\mathcal{T} \in \mathbf{L}[\mathbf{R}^{n\times n}, \mathbf{R}^{n\times n}]$ given by

$$\mathcal{T}(X) = AX + XB \qquad \text{for fixed } A, B \in \mathbf{R}^{n\times n},$$

and let $F : \mathbf{R}^n \to \mathbf{R}^{n\times n}$ be given by

$$F(x) := \operatorname{diag}(x),$$

where the linear transformation $\operatorname{diag}(\cdot) \in \mathbf{L}[\mathbf{R}^n, \mathbf{R}^{n\times n}]$ maps x to the $n \times n$ matrix whose diagonal is x. What is $(\mathcal{T} \circ \operatorname{diag})'(x)$?

Computing Derivatives

Consider the linear transformation $\mathcal{T} \in \mathbf{L}[\mathbf{R}^{n\times n}, \mathbf{R}^{n\times n}]$ given by

$$\mathcal{T}(X) = AX + XB \qquad \text{for fixed } A, B \in \mathbf{R}^{n\times n},$$

and let $F : \mathbf{R}^n \to \mathbf{R}^{n\times n}$ be given by

$$F(x) := \operatorname{diag}(x),$$

where the linear transformation $\operatorname{diag}(\cdot) \in \mathbf{L}[\mathbf{R}^n, \mathbf{R}^{n\times n}]$ maps x to the $n \times n$ matrix whose diagonal is x. What is $(\mathcal{T} \circ \operatorname{diag})'(x)$?

Since both $\mathcal{T}$ and diag are linear, so is $(\mathcal{T} \circ \operatorname{diag})$. Therefore,

$$(\mathcal{T} \circ \operatorname{diag}(\cdot))'(x)(d) = (\mathcal{T} \circ \operatorname{diag}(\cdot))(d) = A\operatorname{diag}(d) + \operatorname{diag}(d)B$$

for all $x \in \mathbf{R}^n$.

Computing Derivatives

Consider the linear transformation $\mathcal{T} \in \mathbf{L}[\mathbf{R}^{n\times n}, \mathbf{R}^{n\times n}]$ given by

$$\mathcal{T}(X) = AX + XB \qquad \text{for fixed } A, B \in \mathbf{R}^{n\times n},$$

and let $F : \mathbf{R}^n \to \mathbf{R}^{n\times n}$ be given by

$$F(x) := \operatorname{diag}(x),$$

where the linear transformation $\operatorname{diag}(\cdot) \in \mathbf{L}[\mathbf{R}^n, \mathbf{R}^{n\times n}]$ maps x to the $n \times n$ matrix whose diagonal is x. What is $(\mathcal{T} \circ \operatorname{diag})'(x)$?

Since both $\mathcal{T}$ and diag are linear, so is $(\mathcal{T} \circ \operatorname{diag})$. Therefore,

$$(\mathcal{T} \circ \operatorname{diag}(\cdot))'(x)(d) = (\mathcal{T} \circ \operatorname{diag}(\cdot))(d) = A \operatorname{diag}(d) + \operatorname{diag}(d) B$$

for all $x \in \mathbf{R}^n$.

What is $\nabla(\mathcal{T} \circ \operatorname{diag}(\cdot))$?

Computing Derivatives

Let $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, and define $f : \mathbf{R}^n \to \mathbf{R}$ by

$$f(x) := \frac{1}{2} \left\| Ax - b \right\|^2 .$$

Compute $\nabla f(x)$ and $\nabla^2 f(x)$.

Computing Derivatives

Let $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, and define $f : \mathbf{R}^n \to \mathbf{R}$ by
$$f(x) := \tfrac{1}{2}\left\|Ax - b\right\|^2 .$$
Compute $\nabla f(x)$ and $\nabla^2 f(x)$.

$$\begin{aligned} f(x+\Delta x) &= \frac{1}{2}\left\|(Ax-b) + A\Delta x\right\|^2 \\ &= \frac{1}{2}\left\|Ax-b\right\|^2 + \langle Ax-b, A\Delta x\rangle + \frac{1}{2}(\Delta x)^T A^T A\Delta x \\ &= f(x) + \langle A^T(Ax-b), \Delta x\rangle + \frac{1}{2}\langle (A^T A)\Delta x, \Delta x\rangle. \end{aligned}$$

Therefore ,

$$\nabla f(x) = A^T(Ax-b) \qquad \text{and} \qquad \nabla^2 f(x) = A^T A.$$

Computing Derivatives

Let $A \in \mathbf{R}^{m\times n}$, $B \in \mathbf{R}^{n\times n}$, and $C \in \mathbf{R}^{n\times k}$, and define $\mathcal{Q} : \mathbf{R}^{n\times n} \times \mathbf{R}^{n\times n} \to \mathbf{R}^{m\times k}$ by

$$\mathcal{Q}(X, Z) = AX^T BZC.$$

$\mathcal{Q}$ is a bilinear mapping in $\mathbf{B}[\mathbf{R}^{n\times n}, \mathbf{R}^{m\times k}]$. This bilinear mapping is a bilinear form if $m = k = 1$, and it is symmetric if $m = k = 1$, $A^T = C$, and $B \in \mathbf{S}^n$.

Compute $\mathcal{Q}'$.

Computing Derivatives

Let $A \in \mathbf{R}^{m\times n}$, $B \in \mathbf{R}^{n\times n}$, and $C \in \mathbf{R}^{n\times k}$, and define $\mathcal{Q} : \mathbf{R}^{n\times n} \times \mathbf{R}^{n\times n} \to \mathbf{R}^{m\times k}$ by

$$\mathcal{Q}(X, Z) = AX^T BZC.$$

$\mathcal{Q}$ is a bilinear mapping in $\mathbf{B}[\mathbf{R}^{n\times n}, \mathbf{R}^{m\times k}]$. This bilinear mapping is a bilinear form if $m = k = 1$, and it is symmetric if $m = k = 1$, $A^T = C$, and $B \in \mathbf{S}^n$.

Compute $\mathcal{Q}'$.

$$\begin{aligned}
&\mathcal{Q}(X + \Delta X, Z + \Delta Z) \\
&\qquad = A(X + \Delta X)^T B(Z + \Delta Z)C \\
&\qquad = AX^T BZC + A(\Delta X)^T BZC + AXB(\Delta Z)C + A(\Delta X)^T B(\Delta Z)C \\
&\qquad = \mathcal{Q}(X, Z) + (A(\Delta X)^T BZC + AXB(\Delta Z)C) + \frac{1}{2}(2\mathcal{Q}(\Delta X, \Delta Z)).
\end{aligned}$$

Hence

$$\mathcal{Q}'(X, Z)(U, V) = \mathcal{Q}(U, Z) + \mathcal{Q}(X, V) \text{ and } \mathcal{Q}''(X, Z)(U, V) = 2\mathcal{Q}(U, V).$$

Computing Derivatives

Let $A \in \mathbf{R}^{m\times n}$, $B \in \mathbf{R}^{n\times n}$, and $C \in \mathbf{R}^{n\times k}$, and define $\mathcal{Q} : \mathbf{R}^{n\times n} \times \mathbf{R}^{n\times n} \to \mathbf{R}^{m\times k}$ by

$$\mathcal{Q}(X, Z) = AX^T BZC.$$

$\mathcal{Q}$ is a bilinear mapping in $\mathbf{B}[\mathbf{R}^{n\times n}, \mathbf{R}^{m\times k}]$. This bilinear mapping is a bilinear form if $m = k = 1$, and it is symmetric if $m = k = 1$, $A^T = C$, and $B \in \mathbf{S}^n$.

Compute $\mathcal{Q}'$.

$$\begin{aligned}
&\mathcal{Q}(X + \Delta X, Z + \Delta Z) \\
&\qquad = A(X + \Delta X)^T B(Z + \Delta Z)C \\
&\qquad = AX^T BZC + A(\Delta X)^T BZC + AXB(\Delta Z)C + A(\Delta X)^T B(\Delta Z)C \\
&\qquad = \mathcal{Q}(X, Z) + (A(\Delta X)^T BZC + AXB(\Delta Z)C) + \frac{1}{2}(2\mathcal{Q}(\Delta X, \Delta Z)).
\end{aligned}$$

Hence

$$\mathcal{Q}'(X, Z)(U, V) = \mathcal{Q}(U, Z) + \mathcal{Q}(X, V) \text{ and } \mathcal{Q}''(X, Z)(U, V) = 2\mathcal{Q}(U, V).$$

Is this true of all bilinear forms regardless of the space?

Computing Derivatives

Let $A \in \mathbf{R}^{m\times n}$, $B \in \mathbf{R}^{n\times n}$, and $C \in \mathbf{R}^{n\times k}$, and define $\mathcal{Q} : \mathbf{R}^{n\times n} \times \mathbf{R}^{n\times n} \to \mathbf{R}^{m\times k}$ by

$$\mathcal{Q}(X, Z) = AX^T BZC.$$

$\mathcal{Q}$ is a bilinear mapping in $\mathbf{B}[\mathbf{R}^{n\times n}, \mathbf{R}^{m\times k}]$. This bilinear mapping is a bilinear form if $m = k = 1$, and it is symmetric if $m = k = 1$, $A^T = C$, and $B \in \mathbf{S}^n$.

Compute $\mathcal{Q}'$.

$$\begin{aligned}
&\mathcal{Q}(X + \Delta X, Z + \Delta Z) \\
&\qquad = A(X + \Delta X)^T B(Z + \Delta Z)C \\
&\qquad = AX^T BZC + A(\Delta X)^T BZC + AXB(\Delta Z)C + A(\Delta X)^T B(\Delta Z)C \\
&\qquad = \mathcal{Q}(X, Z) + (A(\Delta X)^T BZC + AXB(\Delta Z)C) + \frac{1}{2}(2\mathcal{Q}(\Delta X, \Delta Z)).
\end{aligned}$$

Hence

$$\mathcal{Q}'(X, Z)(U, V) = \mathcal{Q}(U, Z) + \mathcal{Q}(X, V) \text{ and } \mathcal{Q}''(X, Z)(U, V) = 2\mathcal{Q}(U, V).$$

Is this true of all bilinear forms regardless of the space?

What is the gradient and Hessian when $m = k = 1$, $A^T = C$, and $B \in \mathbf{S}^n$?

Accuracy of Linear and Quadratic Approximations

Let $U \subset \mathbf{E}$ be open. Consider a function $f\colon U \to \mathbf{R}$ and a point $x \in U$. Multivariate calculus identifies the following two functions as the "best" linear and quadratic approximations of f near x, respectively:

$$l_x(y) := f(x) + \langle \nabla f(x), y - x \rangle,$$

$$Q_x(y) := f(x) + \langle \nabla f(x), y - x \rangle + \tfrac{1}{2}\langle \nabla^2 f(x)(y - x), y - x \rangle.$$

Can we quantify how well these functions approximate f near x?

Accuracy of Linear and Quadratic Approximations

Given $x, y \in \mathbf{E}$ define $\varphi : \mathbf{R} \to \mathbf{R}$ by

$$\varphi(t) := f(x + t(y - x)).$$

Then the following approximation results follow directly from Taylor approximations to φ since $\varphi'(0) = \langle \nabla f(x), y - x \rangle$ and $\varphi''(0) = \langle \nabla^2 f(x)(y - x), y - x \rangle$.

Theorem (Accuracy in approximation)

Consider a C^1-smooth function $f\colon U \to \mathbf{R}$ and two points $x, y \in U$. Then we have

$$f(y) = l_x(y) + \int_0^1 \langle \nabla f(x + t(y - x)) - \nabla f(x), y - x \rangle \, dt.$$

If f is C^2-smooth, then the equation holds:

$$f(y) = Q_x(y) + \int_0^1 \int_0^t \langle (\nabla^2 f(x + s(y - x)) - \nabla^2 f(x))(y - x), y - x \rangle \, ds \, dt.$$

Accuracy of Linear and Quadratic Approximations

Corollary (Accuracy in approximation under Lipschitz conditions)

1 *Suppose that* $f : U \to \mathbf{R}$ *is a* β*-smooth function. Then for any points* $x, y \in U$ *the inequality*

$$\left| f(y) - l_x(y) \right| \leq \frac{\beta}{2} \|y - x\|^2 \quad \textit{holds.}$$

2 *If* f *is* C^2*-smooth and satisfies the estimate*

$$\|\nabla^2 f(y) - \nabla^2 f(x)\|_{\text{op}} \leq M \|y - x\| \qquad \textit{for all } x, y \in U,$$

then the inequality

$$\left| f(y) - Q_x(y) \right| \leq \frac{M}{6} \|y - x\|^3, \qquad \textit{holds for all } x, y \in U.$$

Lipschitz Constants and the Mean Value Theorem

Let $U \subset \mathbf{E}$ be open and $f : U \to \mathbf{R}$ be $\mathcal{C}^1$ on U.

Given $x, y \in U$ with $x \neq y$, set $\varphi(t) := f(x + t(y - x))$.

As we have seen $\varphi'(t) = \langle \nabla f(x + t(y - x)), (y - x) \rangle$. Hence, by the 1-dimensional mean value theorem (MVT), there exists $\bar{t} \in (0, 1)$ such that
$f(y) - f(x) = \varphi(1) - \varphi(0) = \varphi'(\bar{t}) = \langle \nabla f(x + \bar{t}(y - x)), (y - x) \rangle$.

Consequently, given $z \in U$ and $\epsilon > 0$ such that $z + \epsilon\mathbb{B} \subset U$,

$$|f(y) - f(x)| \leq L \, \|y - x\| \qquad \forall \; x, y \in B_\epsilon(z),$$

where

$$L := \max \{ \|\nabla f(v)\| \mid v \in z + \epsilon\mathbb{B} \},$$

and $\mathbb{B} := \{ x \mid \|x\| \leq 1 \}$ is the closed unit ball.

That is, f is locally Lipschitz continuous on U with the local Lipschitz constants given by the gradient. Moreover, if $\operatorname{cl} U$ is compact with ∇f continuous there, then L an be chosen uniformly for all of $\operatorname{cl} U$.

Lipschitz Constants and the Mean Value Theorem

Let $U \subset \mathbf{E}$ be open and $F : U \to \mathbf{R}^m$ be $\mathcal{C}^1$ on U with component functions F_i.

Although, there is no MVT for F, we do have

$$F(y)-F(x) = \int_0^1 \nabla F(x+t(y-x))(y-x)dt = \begin{pmatrix} \int_0^1 \langle \nabla F_1(x+t(y-x)), (y-x)\rangle dt \\ \vdots \\ \int_0^1 \langle \nabla F_m(x+t(y-x)), (y-x)\rangle dt \end{pmatrix}.$$

Hence, given $z \in U$ and $\epsilon > 0$ such that $B_\epsilon(z) \subset U$,

$$\|F(y) - F(x)\| \le L\,\|y - x\| \qquad \forall\ x, y \in B_\epsilon(z),$$

where

$$L := \max\left\{ \|\nabla F(v)\|_{op} \,|\, v \in z + \epsilon\mathbb{B} \right\}.$$

Again, compactness allows us to choose L uniformly on $\operatorname{cl} U$.

First-Order Optimality Conditions

Let $f : \mathbf{E} \to \mathbf{R}$, the directional derivative of f at x in the direction d is given by

$$f'(x; d) := \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}.$$

If f is differentiable at x, then $f'(x; d) = \langle \nabla f(x), d \rangle$.

Theorem (First-order necessary conditions)

Suppose that x is a local minimizer of a function $f : U \to \mathbf{R}$. Then $f'(x; d) \geq 0$ whenever $f'(x; d)$ exists. If f is differentiable at x, then $\nabla f(x) = 0$.

Second-Order Optimality Conditions

Theorem (Second-order conditions)

Consider a $\mathcal{C}^2$-smooth function $f \colon U \to \mathbf{R}$ and fix a point $x \in U$. Then the following are true.

1. *(Necessary conditions) If $x \in U$ is a local minimizer of f, then*

$$\nabla f(x) = 0 \quad \text{and} \quad \nabla^2 f(x) \succeq 0.$$

2. *(Sufficient conditions) If the relations*

$$\nabla f(x) = 0 \quad \text{and} \quad \nabla^2 f(x) \succ 0$$

hold, then x is a local minimizer of f. More precisely, it holds:

$$\liminf_{y \to x} \frac{f(y) - f(x)}{\frac{1}{2}\|y - x\|^2} \geq \lambda_n(\nabla^2 f(x)).$$

Rates of Convergence

Let $\{a_k\} \in \mathbf{R}_+$ be such that $a_k \to 0$.

Sublinear rate: We will say that a_k converges *sublinearly* if there exist constants $c, q > 0$ satisfying

$$a_k \leq \frac{c}{k^q} \qquad \text{for all } k.$$

Larger q and smaller c indicates faster rates of convergence. In particular, given a target precision $\varepsilon > 0$, we have

$$a_k \leq \varepsilon \qquad \forall\ k \geq (\frac{c}{\varepsilon})^{1/q}.$$

The importance of the value of c should not be discounted; the convergence guarantee depends strongly on this value. In applications, it is usually dimension dependent.

Rates of Convergence

Linear rate: The sequence a_k is said to *converge linearly* if there exist constants $c > 0$ and $q \in (0, 1]$ satisfying

$$a_k \leq c \cdot (1-q)^k \qquad \text{for all } k.$$

In this case, we call $(1-q)$ the *linear rate of convergence*. Fix a target accuracy $\varepsilon > 0$, and let us see how large k needs to be to ensure $a_k \leq \varepsilon$. Taking logs we get

$$c \cdot (1-q)^k \leq \varepsilon \quad \Longleftrightarrow \quad k \geq \frac{-1}{\ln(1-q)} \ln\left(\frac{c}{\varepsilon}\right).$$

Taking into account the inequality $\ln(1-q) \leq -q$, we deduce that

$$a_k \leq \varepsilon \qquad \forall\ k \geq \tfrac{1}{q} \ln\left(\frac{c}{\varepsilon}\right).$$

The dependence on q is strong, while the dependence on c is very weak, since the latter appears inside a log.

Rates of Convergence

Quadratic rate: The sequence a_k is said to *converge quadratically* if there is a constant c satisfying

$$a_{k+1} \leq c \cdot a_k^2 \qquad \text{for all } k.$$

The recurrence yields

$$a_{k+1} \leq \frac{1}{c}(ca_0)^{2^{k+1}}.$$

The constant c places conditions on when quadratic convergence begins. In particular, if $ca_0 < 1$, then the inequality $a_k \leq \varepsilon$ holds for all $k \geq \log_2 \ln(\frac{1}{c\varepsilon}) - \log_2(\ln(\frac{1}{ca_0}))$. The dependence on c is negligible.

Note: 2^{-k} converges linearly while 2^{-2^k} converges quadratically.