Appendix A

Definitions

In this chapter, we explain the basic mathematical definitions and concepts that are used in this thesis. Most of the definitions and propositions are taken from [29, 171, 18, 197] that are ready reference used throughout the thesis.

Vector Space: A vector space V is a set that is closed under finite vector-addition and scalar-multiplication. In order for V to be a vector space, the following conditions must hold for all elements $X, Y, Z \in V$ and any scalars r, s:
1. Commutativity ($X+Y=Y+X$)
2. Associativity of vector addition ($(X+Y)+Z = X+(Y+Z)$)
3. Additive identity (For all X, $0+X = X+0=X$)
4. Existence of additive inverse (For any X, there exists a $-X$ such that ($X+(-X) = 0$)
5. Associativity of scalar multiplication ($r(sX)=(rs)X$)
6. Distributivity of scalar sums ($(r+s)X=rX+sX$)
7. Distributivity of vector sums ($r(X+Y)=rX+rY$)
8. Scalar multiplication identity ($1X=X$)

Metric Space: A set S with a global distance function d that, for every two points $x, y \in S$, gives the distance between them as a nonnegative real number $g(x,y)$. A metric space must also satisfy
1. $g(x,y)=0$ iff $x=y$,
2. $g(x,y)=g(y,x)$,
3. The triangle inequality $g(x,y)+g(y,z) >= g(x,z)$.

Cauchy Sequence: A sequence $x_1, x_2, \cdots$ such that the metric $d(x_m, x_n)$ satisfies

$$\lim_{(m,n)\to\infty} d(x_m, x_n) = 0.$$

Complete Metric Space: A metric space in which every Cauchy sequence is convergent.

Norm: A norm $\|\cdot\|$ on $\mathbb{R}^n$ is a function that assigns a scalar $\|x\|$ to every $x \in \mathbb{R}^n$ and that has the following properties:
(a) $\|x\| \geq 0, \ \forall x \in \mathbb{R}^n$.
(b) $\|\alpha x\| = \alpha \cdot \|x\|$ for every scalar α and every $x \in \mathbb{R}^n$.
(c) $\|x\| = 0$ if and only if $x = 0$.
(d) $\|x+y\| \leq \|x\| + \|y\|, \ \forall x, y \in \mathbb{R}^n$ (the triangle inequality).

Banach Spaces: Any convergent sequence in a normed linear space is a Cauchy sequence. However, it may or may not be true in an arbitrary normed linear space that all Cauchy sequences are convergent. A normed linear space X which has the property that all Cauchy sequences are convergent is said to be complete. A complete normed linear space is called a Banach space.

Hilbert Space: A linear space $\mathscr{H}$ with an inner product $\langle \cdot, \cdot \rangle$, which is complete with respect to the norm $\|x\| = \sqrt{\langle x, x \rangle}$ induced by this inner product is called a Hilbert space. Hilbert space is considered to be the generalized Euclidean space with infinite dimensions.

Proposition A.1. *[29] Every bounded and monotonically non-increasing or non-decreasing scalar sequence converges.*

Closed and Open sets: We say that x is a closure point of a subset X of $\mathbb{R}^n$ if there exists a sequence $\{x_n\} \in X$ that converges to x. X is called closed if it is equal to its closure. It is called open if its complement, $\{x \mid x \notin X\}$, is closed. It is called bounded if there exists a scalar c such that $\|x\| \leq c$ for all $x \in X$. It is called compact if it is closed and bounded.

For any $\varepsilon > 0$ and $x^* \in \mathbb{R}^n$,

- $\{x \mid \|x - x^*\| < \varepsilon\}$ is open and is called an open sphere centered at x^*.

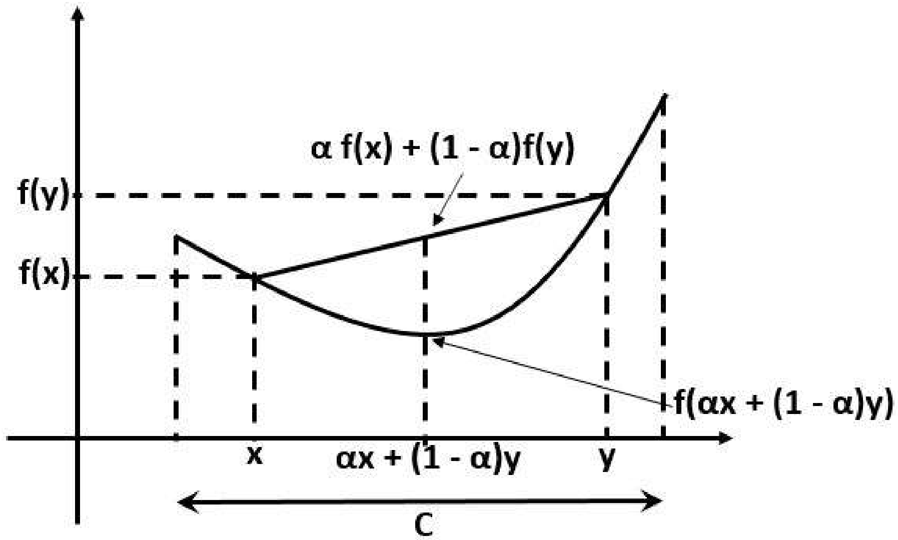

FIGURE A.1: Illustration of convex functions.

- $\{x \mid \|x - x^*\| \leq \varepsilon\}$ is open and is called an open sphere centered at x^*.

Convex Set: A subset C of $\mathbb{R}^n$ is called *convex*, if

$$\alpha x + (1-\alpha)y \in C, \qquad \forall x, y \in C, \forall \alpha \in [0,1].$$

Convex Functions: Let C be a convex subset of $\mathbb{R}^n$. A function $f : C \to \mathbb{R}$ is called convex if

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y), \qquad \forall x, y \in C, \forall \alpha \in [0,1].$$

Figure A.1 demonstrate the convex functions.

Operators: Also called relations or multi-valued functions, an operator T is on $\mathbb{R}^n$ is a subset of $\mathbb{R}^n \times \mathbb{R}^n$.

Subdifferential: The subdifferential relation ∂f of a function $f : \mathbb{R} \to \mathbb{R} \cup \infty$, defined by

$$\partial f = (x, g) | x \in C, \forall z \in \mathbb{R}^n, \ f(z) \geq f(x) + g^T(z - x)$$

The set $\partial f(x)$ is the subdifferential of f at x. Any $g \in \partial f(x)$ is called a subgradient of f at x.

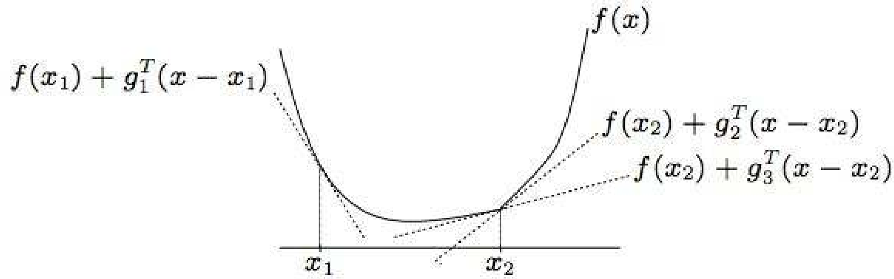

FIGURE A.2: Illustration of Subdifferential [197]

Zeros of an Operator: When $T \ni 0$, we say that x is a zero of T.

Monotone Operators: We say that an operator $T : \mathscr{H} \to \mathscr{H}$ is:

1. monotone if,

$$\langle Tx - Ty, x - y\rangle \geq \eta \|x - y\|^2,$$

for a constant $\eta \geq 0$ and for all $x, y \in X$. If $\eta > 0$, then we say that T is strongly monotone or η-strongly monotone.

2. It is $\delta-$cocoercive if δT is firmly non-expansive, i.e.,

$$\langle Tx - Ty, x - y\rangle \geq \delta \|Tx - Ty\|^2.$$

3. It is an $L-$Lipschitz operator if there exists $L \in [0, \infty)$ such that

$$\|Tx - Ty\| \leq L\|x - y\|, \quad x, y \in \mathscr{H}.$$

It is nonexpansive if it is Lipschitz continuous with constant $L = 1$, i.e.,

$$\|Tx - Ty\| \leq \|x - y\|.$$

It is contraction for $L < 1$.

4. It is a κ-strict pseudo-contraction if for a constant $\kappa \in [0, 1)$,

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 + \kappa\|(I - T)x - (I - T)y\|^2.$$

It is said to be pseudo-contractive if $\kappa = 1$.

Let X be a subset of a Hilbert space $\mathscr{H}$. A mapping $T : X \to X$ is said to be asymptotically nonexpansive if there exists a sequence $\{k_i\}$ of real numbers with $k_i \to 1$ as $i \to \infty$, such that,

$$\|T^i x - T^i y\| \le k_i \|x - y\|, \qquad x, y \in X.$$

Note: Mapping a pair of points by a contraction reduces the distance between them; mapping them by a nonexpansive operator does not increase the distance between them.

Norms: A norm is a function that assigns a strictly positive length or size to each vector in a vector space.

$$\|x\|_p = \begin{cases} (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}, & \text{if } 1 \le p \le \infty \\ \max\ (|x_1|, \cdots, |x_n|), & \text{if } p = \infty \end{cases}$$

Fixed-point Iterations: A fixed point of a function f is a number x^* such that $x^* = f(x^*)$, in other words, it is a solution of the equation $x = f(x)$. The iteration $x_{n+1} \leftarrow f(x_n)$ for $n = 0, 1, \cdots$ is called fixed point iteration.

Note: Assume that $f(x)$ is a continuous function and that $\{x_n\}_{n=0}^{\infty}$ is a sequence generated by fixed point iteration. If $\lim_{n\to\infty} x_n = x^*$, then x^* is a fixed point of $f(x)$.

Weak Convergence: A sequence of points $\{x_n\}$ in a Hilbert space $\mathscr{H}$ is said to converge weakly to a point $x \in \mathscr{H}$ if,

$$\langle x_n, y\rangle \to \langle x, y\rangle, \quad \forall y \in \mathscr{H}.$$

Strong Convergence: A sequence of points $\{x_n\}$ in a Hilbert space $\mathscr{H}$ is said to converge strongly to a point $x \in \mathscr{H}$ if,

$$\|x_n - x\| \to 0 \text{ as } n \to \infty.$$

Maximal Monotone Operator: Operator T is maximal monotone operator, if the graph $G(T)$ is not properly contained in the graph of any other monotone operator. An example of maximal monotone operator is the sub-differential of a convex function.

For any $c > 0$, the resolvent J_c^T defined as $J_c^T = (I + cT)^{-1}$ of a maximal monotone operator T is a non-expansive operator. It is well-known that for any $\lambda > 0$, the resolvent of a maximal monotone operator T is a non-expansive operator [167].

Appendix B

List of Publications

Journals

1. Mridula Verma and K K Shukla (2017) A New Accelerated Proximal Technique for Regression with High-dimensional Datasets. **Knowledge and Information Systems (KAIS)**, doi: 10.1007/s10115-017-1047-z, **(Acceptance Rate $<$ 19.1%)**, **Impact Factor: 2.004**.

2. Mridula Verma, K.K. Shukla (2017) A new accelerated proximal gradient technique for regularized multitask learning framework, Pattern Recognition Letters, Volume 95, 2017, Pages 98-103, ISSN 0167-8655, http://dx.doi.org/10.1016/j.patrec.2017.06.013. **Impact Factor: 1.995**.

3. Mridula Verma, D R Sahu and K K Shukla (2017) VAGA: A Novel Viscosity-based Accelerated Gradient Algorithm: Convergence Analysis and Application to Multitask Regression. **Applied Intelligence**, under minor revision, **Impact Factor: 1.904**.

4. Mridula Verma and K K Shukla (2017) An Extragradient-based Accelerated Algorithm for Microarray Gene Analysis. **Data Mining and Knowledge Discovery**, under review.

5. Mridula Verma and K K Shukla (2017) A New Operator Splitting Algorithm and its Accelerated Variant with Application to Microarray Gene Analysis, **IEEE Transactions on Computational Biology and Bioinformatics**. Communicated.

Conferences

1. Mridula Verma and K K Shukla (2017) Fast Multi-Modal Unified Sparse Representation Learning. Proceedings of 17^{th} **ACM International Conference on Multimedia Retrieval**, June 2017. **(Conference Ranked #1 in the field of Multimedia Retrieval)**, Bucharest, Romania **(acceptance rate: 37%)**.

2. Mridula Verma, Prayas Jain, K K Shukla (2016), A New Faster First Order Iterative Scheme for Sparsity-based Multitask Learning, Proceedings of 2016 **IEEE International Conference on Systems, Man, and Cybernetics (SMC)**, Budapest, Hungary, pp. 1603-1608.

3. Mridula Verma, K K Shukla (2016), Performance Comparison of Proximal Methods for Regression with Nonsmooth Regularizers on Real Datasets, Proceedings of Fifth **International Conference on Computing, Communications and Informatics** (ICACCI-2016), Jaipur. **(Acceptance Rate: 23%)**

4. Mridula Verma, K K Shukla (2016), Efficient Kernel Fuzzy c-means Clustering on Very Large Scale Data using Random Fourier Features, Proceedings of **International Conference on Emerging Research in Computing, Information, Communication and Applications (ERCICA-14)**, Elsevier, Bangalore. **(Acceptance Rate: 33%)**