MTH 995-003: Intro to CS and Big Data Spring 2014

Lecture 6 — Jan 23, 2014

Inst. Mark Iwen *Scribe: Bosu Choi*

Contents

1 Overview

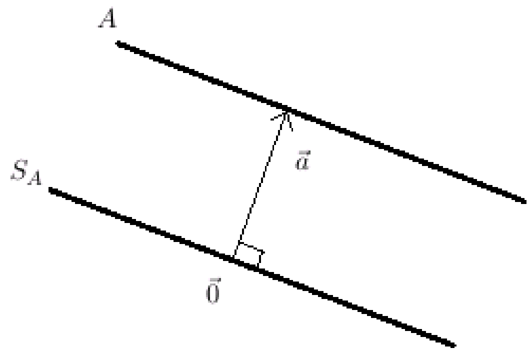
In the last lecture, we discussed feasibility and PCA (i.e., subspace approximation with $\tau = 2$). In this lecture, we discuss subspace approximation with $\tau = \infty$.

2 Subspace Approximation

- In the case of $\tau = \infty$.
- Illustrate using SDP as part of "convex relaxation + rounding" strategy.

2.1 Definitions

Definition 1. *A d-dimensional affine subspace, A, is a set of points* $\{\vec{a} + \vec{x} | \vec{x} \in S_A\}$*, where* S_A *is a d-dimensional subspace and* $\vec{a} \in \mathbb{R}^D$.



Definition 2. *Given a d-dimensional affine subspace, A, we define*

$$\vec{a}_A := argmin_{\vec{x} \in A} \|\vec{x}\|_2.$$

Thus, $A - \vec{a}_A = S_A$ *and* $\vec{a}_A \in S_A^{\perp}$.

Definition 3. *Given an affine subspace A, we can define a projection onto A to be*

$$\Pi_A \vec{x} := \Pi_{S_A} \vec{x} + \vec{a}_A, \forall \vec{x} \in \mathbb{R}^D$$

where Π_{S_A} *is the projection onto* S_A.

Definition 4. *Given a subspace, S,* $\tilde{d} - dimensional$ *for* $\tilde{d} \geq d \geq 1$*, we will define*

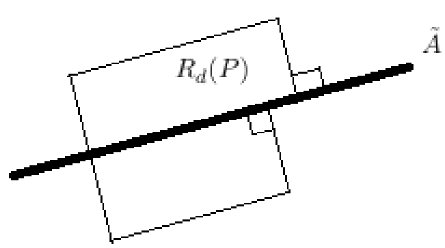
$$\Pi_d(S) = \{all\ d - dimensional\ affine\ subspaces\ of\ S\}.$$

2.2 Main Problem

Given $P = \{\vec{x}_1, \cdots, \vec{x}_N\} \subseteq \mathbb{R}^D$, we want to estimate

$$R_d(P) = inf_{A \in \Pi_d(\mathbb{R}^D)} R_\infty(A, P) := inf_{A \in \Pi_d(\mathbb{R}^D)} \left(max_{j=1,\cdots,N} \|\vec{x}_j - \Pi_A \vec{x}_j\|_2 \right)$$

- How quickly can we find a $\tilde{A} \in \Pi_d(\mathbb{R}^D)$ such that $R_\infty(\tilde{A}, P) \approx R_d(P)$?
- **NOTE** : This is related to bounding box/shape problems in computational geometry.



- **Assumptions about** P:
  1. $\vec{0} \in P$
  2. $\vec{x}_j \in P \Leftrightarrow -\vec{x}_j \in P$

2.3 Solving the Problem

Note that $R_d(P)$ can be found by solving the following optimization problem:

$$R_d^2(P) := min\ \alpha$$

satisfying the constraints:

1. $\sum_{i=1}^{D-d} < \vec{x}_j, \vec{y}_i >^2 \leq \alpha,\ \forall \vec{x}_j \in P$
2. $\|\vec{y}_i\| = 1,\ i = 1, \cdots, D-d.$
3. $< \vec{y}_i, \vec{y}_k >= 0,\ i \neq k$

This problem finds $R_d(P)$ because:

- An optimal d-dimensional subspace A with $R_\infty(A, P) = R_d(P)$ will be given by $(\text{span}\{\vec{y}_1, \cdots, \vec{y}_{D-d}\})^\perp$. That is, we are finding an orthonormal basis for $A^\perp$.
- We are trying to minimize $\|(I - \Pi_A)\vec{x}_j\|_2^2 = \sum_{i=1}^{D-d} < \vec{x}_j, \vec{y}_i >^2$ over all $j = 1, \ldots, N$
- Here, α and the entries of $\vec{y}_1, \cdots, \vec{y}_{D-d} \in \mathbb{R}^D$ are the variables. There are $D(D-d)+1$ total.

2.4 A convex relaxation of the problem [SDP(2)]

Consider this related optimization problem:

Calculate $\tilde{\alpha} :=$ the minimal $\alpha \in \mathbb{R}^+$ satisfying the following constraints for some $Y \in S^D$

1. $$\begin{aligned} \vec{x}_j^T Y \vec{x}_j &= Trace(\vec{x}_j \vec{x}_j^T Y) \leq \alpha, \forall \vec{x}_j \in P. \\ &= \sum_{k=1}^{D} Y_{kk}(x_j)_k^2 + 2\sum_{k=1}^{D} \sum_{l=k+1}^{D} Y_{l,k}(x_j)_l (x_j)_k \leq \alpha, \forall \vec{x}_j \in P. \end{aligned}$$
   $$(\textit{Notice that this is linear in the entries of } Y.)$$
2. $trace(Y) = D - d.$
3. $I - Y \succeq 0.$
4. $Y \succeq 0.$

This problem can be solved as a semidefinite program! Note that:

- The variables are α, and the independent entries of $Y \in S^D$. There are $\frac{D(D+1)}{2} + 1$ total variables.
- Constraint 1 is linear in the variables $\Rightarrow$ it is OK for an SDP.
- Constraint 2 is a linear equality constraint in the variables & so it is OK for an SDP. It implies that the eigenvalues of Y sum to $D - d$.
- Constraint 3 : $I - Y \succeq 0 \Rightarrow I \succeq Y \Rightarrow$ All the eigenvalues of Y are ≤ 1.
- Constraint 4 : All the eigenvalues of Y should be nonnegative.
- Constraints 3 and 4 force all eigenvalues of Y to belong to $[0, 1]$.
- Thus, $\tilde{\alpha}$ can be computed via an SDP.
- **What's left** : Show that $\tilde{\alpha}$ has something to do with $R_d(P)$!

2.5 Homework Problems(due Feb 11th(Tue.))

homework 1. Let $\bar{P} := (P - \vec{x}_1) \cup (\vec{x}_1 - P)$ where $P = \{\vec{x}_1, \cdots, \vec{x}_N\}$
This is now both symmetric about the origin, and contains $\vec{0}$. Prove that $R_d(\bar{P}) \leq 2R_d(P)$.

homework 2. Prove that any affine subspace A with $R_\infty(A, \bar{P}) = R_d(\bar{P})$ will be a subspace (i.e., will have $\vec{a}_A = \vec{0}$).

Problem 3. Show that any optimal orthonormal basis for first optimization problem in section 2.3, $\{\vec{y}_1, \cdots, \vec{y}_{D-d}\}$, satisfies all four constraints for the optimization problem in section 2.4 if we set $Y = \sum_{i=1}^{D-d} \vec{y}_l \vec{y}_l^T$. Conclude that $\tilde{\alpha} \leq R_d^2(P)$.
Hint: The fact that $\tilde{\alpha} \leq R_d^2(P)$ is related to an α you can achieve with this Y in Constraint 1.

2.6 Showing that a solution to [SDP(2)] has something to do with $R_d(P)$

Lemma 1. *Let $\tilde{Y} \in S^D$ be an optimal solution to **[SDP(2)]** in section 2.4. Then $\tilde{Y}$ will have $r \geq D - d$ eigenvalues $\lambda_1, \cdots, \lambda_r \in (0, 1]$ and r(orthogonal unit) eigenvectors $\tilde{v}_1, \cdots, \tilde{v}_r$ with the property that $\sum_{l=1} \lambda_l < \vec{x}_j, \vec{v}_l >^2 \leq R_d^2(P)$, $\forall \vec{x}_j \in P$.*

Proof. Constraints 2 through 4 of [SDP(2)] in section 2.4 guarantee that we have $\lambda_1, \cdots, \lambda_r \in (0, 1]$ for $r \geq D - d$. Also, their associated eigenvectors $\vec{v}_1, \cdots, \vec{v}_r$ are perpendicular. We have that

$$
\begin{aligned}
\sum_{l=1}^{r} \lambda_l < \vec{x}_j, \vec{v}_l >^2 &= Trace\left(\vec{x}_j \vec{x}_j^T \sum_{l=1}^{r} \lambda_l \vec{v}_l \vec{v}_l^T\right) \\
&= Trace(\vec{x}_j \vec{x}_j^T \tilde{Y}) \\
&\leq \tilde{\alpha} \qquad \text{(by [SDP(2)] Constraint 1)} \\
&\leq R_d^2(P) \qquad \text{(by Homework Problem 3)}
\end{aligned}
$$

holds for all $j = 1, \ldots, N$. □