université
PARIS-SACLAY

ÉCOLE
POLYTECHNIQUE
UNIVERSITÉ PARIS-SACLAY

NNT : 2016SACLX031

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
"ECOLE POLYTECHNIQUE"

ECOLE DOCTORALE N° 580
Sciences et technologies de l'information et de la communication (STIC)

Spécialité de doctorat Informatique

Par

Monsieur VU Khac Ky

Projection aléatoire pour l'optimisation de grande dimension

Thèse présentée et soutenue à « LIX, Ecole Polytechnique », le « 5 Juillet 2016 » :

Composition du Jury :

M. PICOULEAU, Christophe	Conservatoire National des Art et Métiers	Président
M. LEDOUX, Michel	University of Toulouse – Paul Sabatier	Rapporteur
M. MEUNIER, Frédéric	Ecole Nationale des Ponts et Chaussées, CERMICS	Rapporteur
M. Walid BEN-AMEUR	Institute TELECOM, TELECOM SudParis, UMR CNRS 5157	Examinateur
M. Frédéric ROUPIN	University Paris 13	Examinateur
Mme. Sourour ELLOUMI	Ecole Nationale Supérieure d'Informatique pour l'industrie et l'Entreprise	Examinateur
M. Leo LIBERTI	LIX, Ecole Polytechnique	Directeur de thèse

Random projections for high-dimensional optimization problems

Vu Khac Ky

A thesis submitted for the degree of
Doctor of Philosophy

Thesis advisor

Prof. Leo Liberti, PhD

Dr. Youssef Hamadi, PhD

Presented to

Laboratoire d'Informatique de l'École Polytechnique (LIX)

University Paris-Saclay

Paris, July 2016

Contents

Acknowledgement

First of all, I would like to express my gratitude to my PhD advisor, Prof. Leo Liberti. He is a wonderful advisor who always encourages, supports and advises me, both in research and career. Under his guidance I have learned about creative ways for solving difficult problems.

I wish to thank Dr. Claudia D'Ambrosio for supervising me when my advisor was in sabbatical years. Her knowledge and professionalism really helped me to build the first steps of my research career. I would also like to acknowledge the supports of my joint PhD advisor, Dr. Youssef Hamadi, during my first 2 years.

I would like to thank Pierre-Louis Poirion for being my long time collaborator. He is a very smart guy who can generate tons of ideas on a problem. It is my pleasure to work with him and his optimism has encouraged me to continue to work on challenging problems.

I would like to thank all my colleagues at the Laboratoire d'informatique de l'École Polytechnique (LIX), including Andrea, Gustavo, Youcef, Claire, Luca, Sonia, and so on... for being my friends and for all their helps.

Last but not least, I would like to thank my wife, Diep, for her unlimited love and support.

This research was supported by a Microsoft Research PhD scholarship.

Résumé

Dans cette thèse, nous allons utiliser les projections aléatoires pour réduire le nombre de variables ou le nombre de contraintes (ou les deux) dans certains problèmes d'optimisation bien connus. En projetant les données dans des espaces de dimension faible, nous obtenons de nouveaux problèmes similaires, mais plus facile à résoudre. De plus, nous essayons d'établir des conditions telles que les deux problèmes (original et projeté) sont fortement liés (en probabilité). Si tel est le cas, alors en résolvant le problème projeté, nous pouvons trouver des solutions approximatives ou une valeur objective approximative pour l'original.

Nous allons appliquer les projections aléatoires pour étudier un certain nombre de problèmes d'optimisation importants, y compris la programmation linéaire et en nombres entiers (Chapitre 2), l'optimisation convexe avec des contraintes linéaires (Chapitre 3), adhésion et rapprocher le plus proche voisins (chapitre 4) et la région-confiance (chapitre sous-problèmes 5). Tous ces résultats sont tirés des documents dont je suis co-auteur avec [26, 25, 24, 27].

Cette thèse sera construite comme suit. Dans le premier chapitre, nous présenterons quelques concepts et des résultats de base en théorie des probabilités. Puisque cette thèse utilise largement les probabilités élémentaire, cette introduction informelle sera plus facile pour les lecteurs avec peu de connaissances sur ce champ pour suivre nos travaux.

Dans le chapitre 2, nous allons présenter brièvement les projections aléatoires et le lemme de Johnson-Lindenstrauss. Nous allons présenter plusieurs constructions des projecteurs aléatoires et expliquer la raison pour laquelle ils fonctionnent. En particulier, les matrices aléatoires sous-gaussienne seront traitées en détail, ainsi que des discussions rapides sur d'autres projections aléatoires.

Dans le chapitre 3, nous étudions les problèmes d'optimisation dans leurs formes de faisabilité. En particulier, nous étudions la soi-disant *problème d'adhésion linéaire restreint*, qui regarde la faisabilité du système $\{Ax = b, x \in \mathcal{C}\}$ où $\mathcal{C}$ est un ensemble qui limite le choix des paramètres x. Cette classe contient de nombreux problèmes importants tels que la faisabilité linéaire et en nombres entiers. Nous proposons d'appliquer une projection aléatoire T aux

contraintes linéaires et d'obtenir le problème projeté correspondant: $\{TAx = Tb, x \in \mathcal{C}\}$. Nous voulons trouver des conditions sur T, de sorte que les deux problèmes de faisabilité sont équivalents avec une forte probabilité. La réponse est simple **quand $\mathcal{C}$ est fini** et borné par un polynôme (en n). Dans ce cas, toute projection aléatoire T avec O $(\log n)$ lignes est suffisante. **Lorsque** $\mathcal{C} = \mathbb{R}^n_+$, nous utilisons l'idée de l'hyperplan séparateur pour séparer b du cône $\{Ax \mid x \geq 0\}$ et montrer que Tb est toujours séparé du cône projeté $\{TAx \mid x \geq 0\}$ sous certaines conditions. Si ces conditions ne tiennent pas, par exemple lorsque le cône $\{Ax \mid x \geq 0\}$ est non-pointue, nous employons l'idée dans le *Johnson-Lindenstrauss lemme* pour prouver que, si $b \notin \{Ax \mid x \geq 0\}$, alors la distance entre b et ce cône est légèrement déformée sous T, reste donc toujours positive. Cependant, le nombre de lignes de T dépend de paramètres inconnus qui sont difficiles à estimer.

Dans le chapitre 4, nous continuons à étudier le problème ci-dessus dans le cas **où $\mathcal{C}$ est un ensemble convexe**. Sous cette hypothèse, on peut définir un cône tangent $\mathcal{K}$ de $\mathcal{C}$ à $x^* \in \arg\min_{x\in\mathcal{C}} \|Ax - b\|$. Nous établissons les relations entre le problème original et le problème projeté sur la base du concept de *largeur gaussienne*, qui est populaire dans les problèmes de *compressed sensing*. En particulier, nous montrons que les deux problèmes sont équivalents avec une forte probabilité pour autant que la projection aléatoire T est échantillonnée à partir des distributions sous-gaussiennes et a au moins $O(\mathbb{W}^2(A\mathcal{K}))$ lignes, où $\mathbb{W}(A\mathcal{K})$ est le largeur Gaussienne de $A\mathcal{K}$. Nous généralisons également ce résultat au cas où T est échantillonnée à partir de systèmes orthonormés randomisés afin d'exploiter leurs propriété de multiplication matrice-vecteur avec des algorithmes plus rapides. Nos résultats sont similaires à ceux de [21], mais ils sont plus utiles dans les applications de préservation de confidentialité, lorsque l'accès aux données originales A, b est limitée ou indisponible.

Dans le chapitre 5, nous étudions le problème **d'adhésion euclidienne:** " Étant donné un vecteur b et un ensemble fermé X dans $\mathbb{R}^n$, décider si $b \in X$ ou pas". Ceci est une généralisation du problème de l'appartenance linéaire restreinte. Nous employons une projection gaussienne aléatoire T pour intégrer à la fois b et X dans un espace de dimension inférieure et étudier la version projetée correspondant: "Décidez si $Tb \in T(X)$ ou non." Lorsque X est fini ou dénombrable, en utilisant un argument simple, nous montrons que les deux problèmes sont équivalents presque sûrement quelle que soit la dimension projetée. Toutefois, ce résultat n'a qu'un intérêt théorique, peut-être en raison d'erreurs d'arrondi dans les opérations à virgule flottante qui rendent difficile son application pratique. Nous abordons cette question en introduisant un seuil $\tau > 0$ et étudier le problème correspondant seuillée: "Décidez si $\mathsf{dist}(Tb, T(X)) \geq \tau$". Dans le cas où X peut être indénombrable, nous montrons que l'original et les projections des problèmes sont également équivalentes si la dimension projetée d est proportionnelle à une dimension intrinsèque de l'ensemble X. En particulier,

nous employons la définition de *dimension doublement* pour prouver que, si $b \notin X$, alors $Sb \notin S(X)$ presque sûrement aussi longtemps que $d = \Omega(\mathsf{DDim}(X))$. ici, $\mathsf{DDim}(X)$ est la dimension de doublement de X, qui est défini comme le plus petit nombre tel que chaque boule dans X peut être couvert par $2^{\mathsf{dd}(X)}$ boules de la moitié du rayon. Nous étendons ce résultat au cas seuillée, et obtenons un plus utile lié pour d. Il se trouve que, en conséquence de ce résultat, nous sommes en mesure pour améliorer une borne de Indyk-Naor sur les plus proches neigbour embeddings Préserver par un facteur de $\frac{\log(1/\delta)}{\varepsilon}$.

Dans le chapitre 6, nous proposons d'appliquer des projections aléatoires pour le problème des **régions de confiance** sous-problème, qui est déclaré comme $\min\{c^\top x + x^\top Q x \mid Ax \leq b,\ \|X\| \leq 1\}$. Ces problèmes se posent dans les méthodes de région de confiance pour faire face à l'optimisation des produits dérivés libre. Soit $P \in \mathbb{R}^{d \times n}$ soit une matrice aléatoire échantillonnée par la distribution gaussienne, nous considérons alors le problème " projeté " suivant:

$$\min\{c^\top P^\top P x + x^\top P^\top R P Q^\top P x \mid AP^\top P x \leq b,\ \|Px\| \leq 1\},$$

qui peut être réduit à $\min\{(Pc)^\top u + u^\top (RPQ^\top) u \mid AP^\top u \leq b,\ \|U\| \leq 1\}$ en définissant $u := Px$. Ce dernier problème est de faible dimension et peut être résolu beaucoup plus rapidement que l'original. Cependant, nous montrons que, si u^* est la solution optimale, alors avec une forte probabilité, $x^* := P^\top u^*$ est une $(1 + O(\varepsilon))$ - approximation pour le problème d'origine. Cela se fait à l'aide de résultats récents de la "concentration de valeurs propres de matrices gaussiennes".

Abstract

In this thesis, we will use random projection to reduce either the number of variables or the number of constraints (or both in some cases) in some well-known optimization problems. By projecting data into lower dimensional spaces, we obtain new problems with similar structures, but much easier to solve. Moreover, we try to establish conditions such that the two problems (original and projected) are strongly related (in probability sense). If it is the case, then by solving the projected problem, we can either find approximate solutions or approximate objective value for the original one.

We will apply random projection to study a number of important optimization problems, including linear and integer programming (Chapter 2), convex optimization with linear constraints (Chapter 3), membership and approximate nearest neighbor (Chapter 4) and trust-region subproblems (Chapter 5). All these results are taken from the papers that I am co-authored with [26, 25, 24, 27].

This thesis will be constructed as follows. In the first chapter, we will present some basic concepts and results in probability theory. Since this thesis extensively uses elementary probability, this informal introduction will make it easier for readers with little background on this field to follow our works.

In Chapter 2, we will briefly introduce to random projection and the Johnson-Lindenstrauss lemma. We will present several constructions of random projectors and explain the reason why they work. In particular, sub-gaussian random matrices will be treated in details, together with some discussion on fast and sparse random projections.

In Chapter 3, we study optimization problems in their feasibility forms. In particular, we study the so-called *restricted linear membership problem*, which asks for the feasibility of the system $\{Ax = b, x \in \mathcal{C}\}$ where $\mathcal{C}$ is some set that restricts the choice of parameters x. This class contains many important problems such as linear and integer feasibility. We propose to apply a random projection T to the linear constraints and obtain the corresponding projected problem: $\{TAx = Tb, x \in \mathcal{C}\}$. We want to find conditions on T, so that the two feasibility

problems are equivalent with high probability. The answer is simple **when $\mathcal{C}$ is finite** and bounded by a polynomial (in n). In that case, any random projection T with $O(\log n)$ rows is sufficient. **When $\mathcal{C} = \mathbb{R}^n_+$**, we use the idea of separating hyperplane to separate b from the cone $\{Ax \mid x \geq 0\}$ and show that Tb is still separated from the projected cone $\{TAx \mid x \geq 0\}$ under certain conditions. If these conditions do not hold, for example when the cone $\{Ax \mid x \geq 0\}$ is non-pointed, we employ the idea in the *Johnson-Lindenstrauss lemma* to prove that, if $b \notin \{Ax \mid x \geq 0\}$, then the distance between b and that cone is slightly distorted under T, thus still remains positive. However, the number of rows of T depends on unknown parameters that are hard to estimate.

In Chapter 4, we continue to study the above problem in the case **when $\mathcal{C}$ is a convex set**. Under that assumption, we can define a tangent cone $\mathcal{K}$ of $\mathcal{C}$ at $x^* \in \arg\min_{x \in \mathcal{C}} \|Ax - b\|$. We establish the relations between the original and projected problems based on the concept of *Gaussian width*, which is popular in *compressed sensing*. In particular, we prove that the two problems are equivalent with high probability as long as the random projection T is sampled from sub-gaussian distributions and has at least $O(\mathbb{W}^2(A\mathcal{K}))$ rows, where $\mathbb{W}(A\mathcal{K})$ is the Gaussian-width of $A\mathcal{K}$. We also extend this result to the case when T is sampled from randomized orthonormal systems in order to exploit its fast matrix-vector multiplication. Our results are similar to those in [21], however they are more useful in privacy-preservation applications when the access to the original data A, b is limited or unavailable.

In Chapter 5, we study the Euclidean **membership problem**: "Given a vector b and a closed set X in $\mathbb{R}^n$, decide whether $b \in X$ or not". This is a generalization of the restricted linear membership problem considered previously. We employ a Gaussian random projection T to embed both b and X into a lower dimension space and study the corresponding projected version: "Decide whether $Tb \in T(X)$ or not". When X is finite or countable, using a straightforward argument, we prove that the two problems are equivalent almost surely regardless the projected dimension. However, this result is only of theoretical interest, possibly due to round-off errors in floating point operations which make its practical application difficult. We address this issue by introducing a threshold $\tau > 0$ and study the corresponding "thresholded" problem: "Decide whether $\mathsf{dist}\,(Tb, T(X)) \geq \tau$". In the case when X may be uncountable, we prove that the original and projected problems are also equivalent if the projected dimension d is proportional to some intrinsic dimension of the set X. In particular, we employ the definition of *doubling dimension* to prove that, if $b \notin X$, then $Sb \notin S(X)$ almost surely as long as $d = \Omega(\mathsf{ddim}(X))$. Here, $\mathsf{ddim}(X)$ is the doubling dimension of X, which is defined as the smallest number such that each ball in X can be covered by at most $2^{\mathsf{dd}(X)}$ balls of half the radius. We extend this result to the thresholded case, and obtain a more useful bound for d. It turns out that, as a consequence of that result, we are able to

improve a bound of Indyk-Naor on the Nearest Neigbour Preserving embeddings by a factor of $\frac{\log(1/\delta)}{\varepsilon}$.

In Chapter 6, we propose to apply random projections for the **trust-region subproblem**, which is stated as $\min\{c^\top x + x^\top Q x \mid Ax \leq b,\ \|x\| \leq 1\}$. These problems arise in trust-region methods for dealing with derivative-free optimization. Let $P \in \mathbb{R}^{d\times n}$ be a random matrix sampled from Gaussian distribution, we then consider the following "projected" problem:

$$\min\{c^\top P^\top Px + x^\top P^\top PQP^\top Px \mid AP^\top Px \leq b,\ \|Px\| \leq 1\},$$

which can be reduced to $\min\{(Pc)^\top u + u^\top (PQP^\top)u \mid AP^\top u \leq b,\ \|u\| \leq 1\}$ by setting $u := Px$. The latter problem is of low dimension and can be solved much faster than the original. However, we prove that, if u^* is its optimal solution, then with high probability, $x^* := P^\top u^*$ is a $(1 + O(\varepsilon))$-approximation for the original problem. This is done by using recent results about the "concentration of eigenvalues" of Gaussian matrices.

Chapter 1

Introduction

Optimization is the process of minimizing or maximizing an objective function over a given domain, which is called the *feasible set*. In this thesis, we consider the following general optimization problem

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to: } & x \in D, \end{aligned}$$

in which $x \in \mathbb{R}^n$, $D \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ is a given function. The feasible set D is often defined by multiple constraints such as: bound constraints ($l \leq x \leq u$), integer constraints ($x \in \mathbb{Z}$ or $\{0, 1\}$), or general constraints ($g(x) \leq O$ for some $g : \mathbb{R}^n \to \mathbb{R}^m$).

In the digitization age, data becomes cheap and easy to obtain. That results in many new optimization problems with extremely large sizes. In particular, for the same kind of problems, the numbers of variables and constraints are huge. Moreover, in many application settings such as those in *Machine Learning*, an accurate solution is less preferred as approximate but robust ones. It is a real challenge for traditional algorithms, which are used to work well with average-size problems, to deal with these new circumstances.

Instead of developing algorithms that scale up well to solve these problems directly, one natural idea is to transform them into small-size problems that strongly relates to the originals. Since the new ones are of manageable sizes, they can still be solved efficiently by classical methods. The solutions obtained by these new problems, however, will provide us insight into the original problems. In this thesis, we will exploit the above idea to solve some high-dimensional optimization problems. In particular, we **apply a special technique called *random projection* to embed the problem data into low dimensional spaces, and approximately reformulate the problem in such a way that it becomes very easy to solve but still captures the most important information**.

1.1 Random projection versus Principal Component Analysis

Random projection (defined formally in Chapter 2) is the process of mapping high-dimensional vectors to a lower-dimensional space by a random matrix. Examples of random projectors are matrices with i.i.d Gaussian or Rademacher entries. These matrices are constructed in certain ways, so that, with high probability, they can well-approximate many geometrical structures such as distances, inner products, volumes, curves, ... The most interesting feature is that, they are often very "fat" matrices, i.e. the number of rows is significantly smaller than the number of columns. Therefore, they can be used as a dimension reduction tool, simply by taking matrix-vector multiplications.

Despite of its simplicity, random projection works very well and comparable to many other classical dimension reduction methods. One method that is often compared to random projection is the so-called *Principal Component Analysis* (PCA). PCA attempts to find a set of orthonormal vectors $\{\xi_1, \xi_2, \ldots\}$ that better represent the data points. These vectors are often called principal components. In particular, the first component ξ_1 is found as the vector with the largest variance, i.e.

$$\xi_1 = \arg\max_{\|u\|=1} \sum_{i=1}^{n} \langle x_i, u\rangle^2,$$

and inductively, ξ_i is found as the vector with the largest variance among all the vectors that are orthogonal to $\xi_1, \ldots, \xi_{i-1}$. In order to apply PCA for dimension reduction, we simply take k first components out to obtain a matrix $\Xi_k = (\xi_1 \ldots \xi_k)$, and then form the new (lower-dimensional) data point $T_k = X\Xi_k$.

Note that, PCA closely relates to the singular vector decomposition (SVD) of the matrix X. Recall that, any matrix X can be written in the SVD form as the product of $U\Sigma V^\top$, in which U, V are $(n \times n)$-orthogonal matrices (i.e. $UU^\top = VV^\top = I_n$) and Σ is a diagonal matrix with nonnegative entries that are ordered in the decreasing way. The matrix T_k (discussed previously) can now be written as $T_k = U_k\Sigma_k$, by truncating the singular values in that decomposition.

It is easy to see that, as opposed to PCA and SVD, random projection is much cheaper to compute. The complexity of constructing a random projector is often fractional to the number of entries, i.e. $O(nm)$, and that is significantly smaller than the complexity $O(nm^2 + m^3)$ of PCA. Moreover, random projections are data-independent, i.e. they are always constructed the same way regardless how the point set is distributed. This property is often called *oblivious*, and it is one of the main advantages of random projection over other dimension reduction techniques. In many applications, the number of data points are often very large

and/or might not be known at hand (in the case of online and stream computations). In these circumstances, it is expensive and even impossible to exploit the information of the data point to construct principal components in the PCA method. Random projection, therefore, is the only choice.

1.2 Structure of the thesis

In this thesis, we will use random projection to reduce either the number of variables or the number of constraints (or both in some cases) in some well-known optimization problems. By projecting data into lower dimensional spaces, we obtain new problems with similar structures, but much easier to solve. Moreover, we try to establish conditions such that the two problems (original and projected) are strongly related (in probability sense). If it is the case, then by solving the projected problem, we can either find approximate solutions or approximate objective value for the original one.

We will apply random projection to study a number of important optimization problems, including linear and integer programming (Chapter 2), convex optimization with linear constraints (Chapter 3), membership and approximate nearest neighbor (Chapter 4) and trust-region subproblems (Chapter 5). All these results are taken from the papers that I am co-authored with [26, 25, 24, 27].

The rest of this thesis will be constructed as follows. At the end of this chapter, we will present some basic concepts and results in probability theory. Since this thesis extensively uses elementary probability, this informal introduction will make it easier for readers with little background on this field to follow our works.

In Chapter 2, we will briefly introduce to random projection and the Johnson-Lindenstrauss lemma. We will present several constructions of random projectors and explain the reason why they work. In particular, sub-gaussian random matrices will be treated in details, together with some discussion on fast and sparse random projections.

In Chapter 3, we study optimization problems in their feasibility forms. In particular, we study the so-called *restricted linear membership problem*, which asks for the feasibility of the system $\{Ax = b, x \in \mathcal{C}\}$ where $\mathcal{C}$ is some set that restricts the choice of parameters x. This class contains many important problems such as linear and integer feasibility. We propose to apply a random projection T to the linear constraints and obtain the corresponding projected problem: $\{TAx = Tb, x \in \mathcal{C}\}$. We want to find conditions on T, so that the two feasibility problems are equivalent with high probability. The answer is simple **when $\mathcal{C}$ is finite** and bounded by a polynomial (in n). In that case, any random projection T with $O(\log n)$

rows is sufficient. **When $\mathcal{C} = \mathbb{R}^n_+$**, we use the idea of separating hyperplane to separate b from the cone $\{Ax \mid x \geq 0\}$ and show that Tb is still separated from the projected cone $\{TAx \mid x \geq 0\}$ under certain conditions. If these conditions do not hold, for example when the cone $\{Ax \mid x \geq 0\}$ is non-pointed, we employ the idea in the *Johnson-Lindenstrauss lemma* to prove that, if $b \notin \{Ax \mid x \geq 0\}$, then the distance between b and that cone is slightly distorted under T, thus still remains positive. However, the number of rows of T depends on unknown parameters that are hard to estimate.

In Chapter 4, we continue to study the above problem in the case **when $\mathcal{C}$ is a convex set**. Under that assumption, we can define a tangent cone $\mathcal{K}$ of $\mathcal{C}$ at $x^* \in \arg\min_{x \in \mathcal{C}} \|Ax - b\|$. We establish the relations between the original and projected problems based on the concept of *Gaussian width*, which is popular in *compressed sensing*. In particular, we prove that the two problems are equivalent with high probability as long as the random projection T is sampled from sub-gaussian distributions and has at least $O(\mathbb{W}^2(A\mathcal{K}))$ rows, where $\mathbb{W}(A\mathcal{K})$ is the Gaussian-width of $A\mathcal{K}$. We also extend this result to the case when T is sampled from randomized orthonormal systems in order to exploit its fast matrix-vector multiplication. Our results are similar to those in [21], however they are more useful in privacy-preservation applications when the access to the original data A, b is limited or unavailable.

In Chapter 5, we study the Euclidean **membership problem**: "Given a vector b and a closed set X in $\mathbb{R}^n$, decide whether $b \in X$ or not". This is a generalization of the restricted linear membership problem considered previously. We employ a Gaussian random projection T to embed both b and X into a lower dimension space and study the corresponding projected version: "Decide whether $Tb \in T(X)$ or not". When X is finite or countable, using a straightforward argument, we prove that the two problems are equivalent almost surely regardless the projected dimension. However, this result is only of theoretical interest, possibly due to round-off errors in floating point operations which make its practical application difficult. We address this issue by introducing a threshold $\tau > 0$ and study the corresponding "thresholded" problem: "Decide whether $\mathsf{dist}\,(Tb, T(X)) \geq \tau$". In the case when X may be uncountable, we prove that the original and projected problems are also equivalent if the projected dimension d is proportional to some intrinsic dimension of the set X. In particular, we employ the definition of *doubling dimension* to prove that, if $b \notin X$, then $Sb \notin S(X)$ almost surely as long as $d = \Omega(\mathsf{ddim}(X))$. Here, $\mathsf{ddim}(X)$ is the doubling dimension of X, which is defined as the smallest number such that each ball in X can be covered by at most $2^{\mathsf{dd}(X)}$ balls of half the radius. We extend this result to the thresholded case, and obtain a more useful bound for d. It turns out that, as a consequence of that result, we are able to improve a bound of Indyk-Naor on the Nearest Neigbour Preserving embeddings by a factor of $\frac{\log(1/\delta)}{\varepsilon}$.

In Chapter 6, we propose to apply random projections for the **trust-region subproblem**, which is stated as $\min\{c^\top x + x^\top Q x \mid Ax \leq b,\ \|x\| \leq 1\}$. These problems arise in trust-region methods for dealing with derivative-free optimization. Let $P \in \mathbb{R}^{d\times n}$ be a random matrix sampled from Gaussian distribution, we then consider the following "projected" problem:

$$\min\{c^\top P^\top Px + x^\top P^\top PQP^\top Px \mid AP^\top Px \leq b,\ \|Px\| \leq 1\},$$

which can be reduced to $\min\{(Pc)^\top u + u^\top (PQP^\top)u \mid AP^\top u \leq b,\ \|u\| \leq 1\}$ by setting $u := Px$. The latter problem is of low dimension and can be solved much faster than the original. However, we prove that, if u^* is its optimal solution, then with high probability, $x^* := P^\top u^*$ is a $(1 + O(\varepsilon))$-approximation for the original problem. This is done by using recent results about the "concentration of eigenvalues" of Gaussian matrices.

1.3 Preliminaries on Probability Theory

A **probability space** is mathematically defined as a triple $(\Omega, \mathcal{A}, \mathbb{P})$, in which

- Ω is a non-empty set (**sample space**)
- $\mathcal{A} \subseteq 2^\Omega$ is a σ-algebra over Ω (**set of events**) and
- $\mathbb{P}$ is a **probability measure** on $\mathcal{A}$.

A family $\emptyset \in \mathcal{A}$ of subsets of Ω is called a **σ-algebra** (over Ω) if it is closed under compliments and countable unions. More precisely,

- $\emptyset, \Omega \in \mathcal{A}$. is a non-empty set (**sample space**)
- If $E \in \mathcal{A}$ then $E^c := \Omega \setminus E \in \mathcal{A}$.
- If $E_1, E_2, \ldots \in \mathcal{A}$ then $\bigcup_{i=1}^\infty E_i \in \mathcal{A}$.

A function $\mathbb{P} : \mathcal{A} \to [0, 1]$ is called a **probability measure** if it is countably additive and its value over the entire sample space is equal to one. More precisely,

- If $A_1, A_2, \ldots$ are a countable collection of pairwise disjoint sets, then

$$\mathbb{P}(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mathbb{P}(A_i),$$

- $\mathbb{P}(\Omega) = 1$.

Each $E \in \mathcal{A}$ is called an **event** and $\mathbb{P}(E)$ is called the probability that the event E occurs.

If $E \in \mathcal{A}$ and $\mathbb{P}(E) = 1$, then E is called an **almost sure** event.

A function $X : \mathcal{A} \to \mathbb{R}^n$ is called a **random variable** if for all Borel set Y in $\mathbb{R}$, $X^{-1}(Y) \in \mathcal{A}$.

Given a random variable X, the **distribution function** of X, denoted by F_X is defined as follows:

$$F_X(x) := \mathbb{P}(\omega : X(\omega) \leq x) = \mathbb{P}(X \leq x)$$

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

Given a random variable X, the **density function** of X is any measurable function with the property that

$$\mathbb{P}[X \in A] = \int_{X^{-1}A} dP = \int_A f \, d\mu$$

for any $A \in \mathcal{A}$.

The following distribution functions are used in this thesis:

- Discrete distribution: X only takes values $x_1, x_2 \ldots$, each with probability $p_1, p_2, \ldots$ where $0 \leq p_i$ and $\sum_i p_i = 1$.
- Rademacher distribution: X only takes values -1 and 1, each with probability $\frac{1}{2}$.
- Uniform distribution: X takes values in the interval $[a, b]$ and has the density function

$$f(x) = \frac{1}{b-a}.$$

- Normal distribution: X has the density function

$$f(x) = \frac{1}{\sqrt{2\pi}} \, e^{-x^2/2}.$$

The **expectation** of a random variable X is defined as follows:

- $\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i p_i$ if X has discrete distribution: $\mathbb{P}(X = x_i) = p_i$ for $i = 1, 2, \ldots$.
- $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx$, if X has a (continuous) density function $f(.)$.

The **variance** of a random variable X is defined as follows:

- $\mathrm{Var}(X) = \sum_{i=1}^{\infty} x_i^2 p_i$ if X has discrete distribution: $\mathbb{P}(X = x_i) = p_i$ for $i = 1, 2, \ldots$.
- $\mathrm{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x) dx$, if X has a (continuous) density function $f(.)$.

The following property, which is called **union bound**, will be used very often in this thesis:

Lemma 1.3.1 (Union bound)**.** *Given events $A_1, A_2, \ldots$ and positive numbers $\delta_1, \delta_2, \ldots$ such that for each i, the event A_i occurs with probability at least $1 - \delta_i$. Then the probability that all these events occur is at least $1 - \sum_{i=1}^{\infty} \delta_i$.*

We also use the following simple but very useful inequality:
Markov inequality: For any nonnegative random variable X and any $t > 0$, we have

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

Note that, if we have a nonnegative increasing function f, then we can apply Markov inequality to obtain

$$\mathbb{P}(X \geq t) = \mathbb{P}(f(X) \geq f(t)) \leq \frac{\mathbb{E}(f(X))}{f(t)}.$$

Chapter 2

Random projections and Johnson-Lindenstrauss lemma

2.1 Johnson-Lindenstrauss lemma

One of the main motivations for the development of random projections is the so-called *Johnson-Lindenstrauss lemma* (JLL), which is found by William B. Johnson and Joram Lindenstrauss in their 1984 seminal paper [14]. The lemma asserts that any subset of $\mathbb{R}^m$ can be embedded into a low-dimension space $\mathbb{R}^k$ ($k \ll m$), whilst keeping the Euclidean distances between any two points of the set almost the same. Formally, it is stated as follows:

Theorem 2.1.1 (Johnson-Lindenstrauss Lemma [14])**.** *Given $\varepsilon \in (0,1)$ and $A = \{a_1, \ldots, a_n\}$ be a set of n points in $\mathbb{R}^m$. Then there exists a mapping $T : \mathbb{R}^m \to \mathbb{R}^k$, where $k = O(\varepsilon^{-2} \log n)$, such that*

$$(1-\varepsilon)\|a_i - a_j\|^2 \leq \|T(a_i) - T(a_j)\|^2 \leq (1+\varepsilon)\|a_i - a_j\|^2 \tag{2.1}$$

for all $1 \leq i, j \leq n$.

To see why this theorem is important, let's imagine we have a billion of points in $\mathbb{R}^m$. According to JLL, we can compress these points by projecting them into $\mathbb{R}^k$, with $k = O(\varepsilon^{-2} \log n) \approx O(20\,\varepsilon^{-2})$. For reasonable choices of the error ε, the projected dimension k might be much smaller than m (for example, when $m = 10^6$ and $\varepsilon = 0.01$). The effect is even more significant for larger instances, mostly due to the rapid decrease of the log-function.

Note that in JLL, the magnitude of the projected dimension k only depends on the number of data points n and a predetermined error ε, but not on the original dimension m. Therefore, JLL is more meaningful for the "big-data" cases, i.e. when m and n are huge. In contrast, if

m is small, then small values of ε will result in dimensions k that are larger than m. In that case the applying of JLL is not useful.

The existence of the map T in JLL is shown by probabilistic methods. In particular, T is drawn from some well-structured classes of random maps $\mathcal{T}$, in such a way one can prove that the inequalities (2.1) hold for all $i \leq i, j \leq n$ with some positive probability. This can be done if the random map $\mathcal{T}$ satisfies for all $x \in \mathbb{R}^m$:

$$\mathbb{P}\left[(1-\varepsilon)\|x\|^2 \leq \|\mathcal{T}(x)\|^2 \leq (1+\varepsilon)\|x\|^2\right] > 1 - \frac{2}{(n-1)n}. \tag{2.2}$$

Indeed, if it is the case, we have

$$\mathbb{P}\left[(1-\varepsilon)\|x_i - x_j\|^2 \leq \|\mathcal{T}(x_i - x_j)\|^2 \leq (1+\varepsilon)\|x_i - x_j\|^2\right] > 1 - \frac{2}{(n-1)n},$$

for all $1 \leq i < j \leq n$. It is only left to apply the union bound for $\binom{n}{2}$ pairs of points (x_i, x_j).

There are several ways to construct a random map $\mathcal{T}$ satisfying the requirement (2.2). In the original paper of Johnson and Lindenstrauss ([14]), $\mathcal{T}$ is constructed as the orthogonal projection on a k-dimensional random subspace of $\mathbb{R}^m$. Later on, P. Indyk and R. Motwani [13] noticed that $\mathcal{T}$ can be defined simply as a random matrix whose entries are i.i.d Gaussian random variables. Another breakthrough is made by Achlioptas [1], in which the entries of $\mathcal{T}$ are greatly simplified to i.i.d random variables of ± 1 values, each with probability $\frac{1}{2}$. He also gave another interesting construction

$$\mathcal{T}_{i,j} = \begin{cases} +1 & \text{with probability } \frac{1}{6}, \\ 0 & \text{with probability } \frac{2}{3}, \\ -1 & \text{with probability } \frac{1}{6}. \end{cases}$$

This construction is particularly useful because it only requires $\frac{1}{3}$ of the entries to be non-zero, therefore the matrix $\mathcal{T}$ becomes relatively sparse. The sparsity of $\mathcal{T}$ leads to faster matrix-vector multiplications, which are useful in many applications.

After the discoveries of these results, many researchers continue to find more sophisticated constructions of the random map $\mathcal{T}$ that are suitable for specific problems. These researches can be divided into 2 main branches: faster constructions (i.e. to obtain fast matrix-vector multiplications) and sparser constructions (i.e. to obtain sparse random matrices). As opposed to the common intuition, many fastest random matrices are dense. However, sparse matrices are obvious relatively "fast".

In the next section, we will give a formal definition of random projections. Several popular constructions of random projections will be briefly introduced in Section 2.3.

For the simplicity of notations, for the rest of the thesis, we will use T (instead of $\mathcal{T}$) to denote a random map. Moreover, since we will only work with linear maps, T is also treated as a random matrix. Thus, the expression $T(x)$ is best written as the matrix-vector product Tx and T_{ij} will stand for the $(i,j)-$entry of T. The terminologies "random projection", "random mapping" and "random matrix", therefore can be used exchangeable.

2.2 Definition of random projections

2.2.1 Normalized random projections

It might be useful and easy to follow if we give a formal definition of random projections (RP). However, there is no general agreement on what an RP is. Naturally, we first provide the properties that are desired and look for structures that follow them. Since we will focus on the applications of existing random projections instead of constructing new ones, for convenience, we will give the following definition (motivated by the unpublished manuscript of Jiří Matoušek [19]). It contains the property that we will mostly deal with in this thesis.

Definition 2.2.1. *A random linear map $T : \mathbb{R}^m \to \mathbb{R}^k$ is called a* random projection *(or random matrix) if for all $\varepsilon \in (0,1)$ and all vectors $x \in \mathbb{R}^m$, we have:*

$$\mathbb{P}(\,(1-\varepsilon)\|x\|^2 \leq \|T(x)\|^2 \leq (1+\varepsilon)\|x\|^2\,) \geq 1 - 2e^{-\mathcal{C}\varepsilon^2 k} \tag{2.3}$$

for some universal constant $\mathcal{C} > 0$ (independent of m, k, ε).

It should be noted that, given an RP, one can show the existence of a map T satisfying conditions in the Johnson-Lindenstrauss lemma. Indeed, to obtain a positive probability, it is sufficient to choose k such that $1 - 2e^{-\mathcal{C}\varepsilon^2 k} \geq 1 - \frac{2}{(n-1)n}$, and this can be done with any $k > (\frac{2}{\mathcal{C}})\frac{\log n}{\varepsilon^2}$. More interesting, the probability that we can successfully find such a map (by sampling) is very high. For example, if we want this probability to be at least, say 99.9%, by the union bound, we can simply choose any k such that

$$1 - n(n-1)e^{-\mathcal{C}\varepsilon^2 k} > 1 - \frac{1}{1000}.$$

This means k can be chosen to be $k = \lceil \frac{\ln(1000)+2\ln(n)}{\mathcal{C}\varepsilon^2} \rceil \leq \lceil \frac{7+2\ln(n)}{\mathcal{C}\varepsilon^2} \rceil$. Therefore, by slightly increasing the projected dimension, we almost always obtain a "good" mapping without having to re-sample.

2.2.2 Preservation of scalar products

From the definition, we can immediately see that an RP also preserves the scalar product with high probability. Indeed, given any $x, y \in \mathbb{R}^n$, then by applying the definition of RP on two vectors $x+y$, $x-y$ and using the union bound, we have

$$\begin{aligned}|\langle Tx, Ty\rangle - \langle x, y\rangle| &= \frac{1}{4}\big|\|T(x+y)\|^2 - \|T(x-y)\|^2 - \|x+y\|^2 + \|x-y\|^2\big| \\ &\leq \frac{1}{4}\big|\|T(x+y)\|^2 - \|x+y\|^2\big| + \frac{1}{4}\big|\|T(x-y)\|^2 - \|x-y\|^2\big| \\ &\leq \frac{\varepsilon}{4}(\|x+y\|^2 + \|x-y\|^2) = \frac{\varepsilon}{2}(\|x\|^2 + \|y\|^2),\end{aligned}$$

with probability at least $1 - 4e^{-\mathcal{C}\varepsilon^2 k}$. We can actually strengthen it to obtain the following useful lemma:

Lemma 2.2.2. *Let $T : \mathbb{R}^m \to \mathbb{R}^k$ be a random projection and $0 < \varepsilon < 1$. Then there is a universal constant $\mathcal{C}$ such that, for any $x, y \in \mathbb{R}^n$:*

$$\langle Tx, Ty\rangle = \langle x, y\rangle \pm \varepsilon\|x\|.\|y\|$$

with probability at least $1 - 4e^{-\mathcal{C}\varepsilon^2 k}$.

Proof. Apply the above result for $u = \frac{x}{\|x\|}$ and $v = \frac{y}{\|y\|}$. □

2.2.3 Lower bounds for projected dimension

It is interesting to know whether we can obtain a lower value for the dimension k if we use a smarter construction of a map T (such as nonlinear ones). It turns out that the answer is negative, i.e. the value $k = O(\varepsilon^{-2}\log n)$ is almost optimal. Indeed, Noga Alon shows in ([5]) that there exists a set of n points such that the dimension k has to be at least $k = O(\frac{\log n}{\varepsilon^2 \log(1/\varepsilon)})$ in order to preserve the distances between all pairs of points. Moreover, when the mapping T is required to be linear, then Larsen and Nelson [16]) are able to prove that $k = O(\varepsilon^{-2}\log n)$ is actually the best possible. Therefore, the *linearity* requirement of T in the definition 2.2.1 is quite natural.

2.3 Constructions of random projections

In this section, we introduce several popular constructions of random projections. We first consider sub-gaussian random matrices, the simplest case that contains many well-known constructions that are mentioned in the previous section. Then we move to fast constructions

in the subsection 2.3.2. For simplicity, we will discard the scaling factor $\sqrt{\frac{n}{k}}$ in the random matrix T. This does not affect the main ideas we present but makes them clearer and more concise.

2.3.1 Sub-gaussian random projections

Perhaps the simplest construction of a random projection is matrices with i.i.d entries which are drawn from a certain "good distribution". Some of good distributions are: Normal $\mathcal{N}(0,1)$ (by Indyk and Motwani ([13]), Rademacher (± 1) (by Achlioptas [1]), Uniform $U(-a,a)$, ... In [18], Matoušek shows that all these distributions turn out to be special cases of a general class of the so-called *sub-Gaussian* distributions. He proves that an RP can still be obtained by sampling under this general distribution.

Definition 2.3.1. *A random variable X is called sub-Gaussian (or to has a sub-Gaussian tail) if there are constants C, δ such that for any $t > 0$:*

$$\mathbb{P}(|X| > t) \leq Ce^{-\delta t^2}.$$

A random variable X is called sub-Gaussian up to t_0 if there are constants C, δ such that for any $t > t_0$:

$$\mathbb{P}(|X| > t) \leq Ce^{-\delta t^2}.$$

As the name suggests, this family of distributions strongly relates to Gaussian distribution. Intuitively, a sub-Gaussian random variable has a strong tail decay property, similar to that of a Gaussian distribution. One of its useful properties is that, a linear combination of sub-Gaussian random variables (of the uniform constants C, δ) is again sub-Gaussian. Moreover,

Lemma 2.3.2. *Let X be a random variable with $\mathbb{E}(X) = 0$. If $\mathbb{E}(e^{uX}) \leq e^{Cu^2}$ for some constant C and any $u > 0$, then X has a sub-gaussian tail. In reverse, if $Var(X) = 1$ and X has a sub-gaussian tail, then $\mathbb{E}(e^{uX}) \leq e^{Cu^2}$ for all $u > 0$ and some constant C.*

The following lemma states that, the sum of squares of sub-Gaussian random projections also has a sub-Gaussian tail up to a constant.

Lemma 2.3.3 (Matoušek [18])**.** *Let $Y_1, \ldots, Y_k$ be independent random variables with $\mathbb{E}(Y_j) = 0$ and $Var(E) = 1$ and a uniform sub-Gaussian tail. Then*

$$Z = \frac{1}{\sqrt{k}}(Y_1^2 + \ldots + Y_k^2 - k)$$

has a sub-Gaussian tail up to $\sqrt{k}$.

Now let $T \in \mathbb{R}^{k\times m}$ be a matrix whose entries are random variables with expectation 0, variance $\frac{1}{k}$ and a uniform sub-Gaussian tail. Then, for any $x \in \mathbb{R}^m$,

$$\|Tx\|^2 - 1 = (T_1 x)^2 + \ldots + (T_k x)^2 - 1$$

has the distribution the same as the variable $\frac{1}{\sqrt{k}}Z$ in the above lemma. By the definition of sub-Gaussian tail, we have

$$\text{Prob } (|\|Tx\|^2 - 1| \geq \varepsilon) = \text{Prob } (|Z| \geq \varepsilon\sqrt{k}) \leq Ce^{-\delta\varepsilon^2 k},$$

which then implies that T is an RP.

2.3.2 Fast Johnson-Lindenstrauss transforms

In many applications, the bottle-neck in applying random projection techniques is the cost of matrix-vector multiplications. Indeed, the complexity of multiplying a $k \times n$ matrix T to a vector is of order $O(kn)$. Even with Achlioptas' sparse construction, the computation time is only decreased by the factor of 3. Therefore, it is important to construct the random matrix T such that the operations Tx can be done as fast as possible.

It is natural to expect that, the sparser the matrix T we can construct, the faster the product Tx becomes. However, due to the *uncertainty principle* in analysis, if the vector x is also sparse, then its image under a sparse matrix T can be largely distorted. Therefore, a random projection that satisfies Johnson-Lindenstrauss lemma cannot be too sparse.

One of the ingenious ideas to construct fast random projectors is given by Ailon and Chazelle [3], in which they propose the so-called Fast Johnson Lindenstrauss Transform (FJLT). The idea is to precondition a vector (possibly sparse) by an orthogonal matrix, in order to enlarge its support. After the precondition step, we obtain a "smooth" vector, which can now be projected under a sparse random projector. More precisely, a FJLT is constructed as a product of three real-valued matrices $T = PHD$, which are defined as follows:

- P is a $k \times d$ matrix whose elements are independently distributed as follows:

$$P_{ij} = \begin{cases} 0 & \text{with probability } 1-q \\ \sim \mathcal{N}(0, \frac{1}{q}) & \text{with probability } q. \end{cases}$$

  Here q is a sparsity constant given by $q = \min\{\Theta(\frac{\varepsilon^{p-2}\log^p(n)}{d}), 1\}$

- H is a $d \times d$ normalized Walsh–Hadamard matrix:

$$H_{ij} = \frac{1}{\sqrt{d}}(-1)^{\langle i-1, j-1\rangle}$$

  where $\langle i, j\rangle$ is the dot-product of the m-bit vectors i, j expressed in binary.

- D is a $d \times d$ diagonal matrix, where each D_{ii} independently takes value equal to -1 or 1, each with probability $\frac{1}{2}$.

Note that, in the above definition, the matrices P and D are random and H is deterministic. Moreover, both H and D are orthogonal matrices, therefore it is only expected that P obeys the low-distortion property, i.e. $\|Py\|$ is not so different from y. However, the vectors y being considered are not the entire set of unit vectors, but are restricted to only those that have the form HDx. The two matrices H, D play the role of "smoothening" x, so that we have the following property:

Property: Given a set X of n unit vectors. Then

$$\max_{x \in X} \|HDx\|_\infty = O(\frac{\log n}{\sqrt{k}}),$$

with probability at least $1 - \frac{1}{20}$.

The main theorem regarding FJLT is stated as follows:

Theorem 2.3.4 ([3])**.** *Given a set X of n unit vectors in $\mathbb{R}^n$, $\varepsilon < 1$, and $p \in \{1, 2\}$. Let T be a FJLT defined as above. Then with probability at least 2/3, the following two events occur:*

1. *For all $x \in X$:*
$$(1 - \varepsilon)\alpha_p \leq \|Tx\|_p \leq (1 + \varepsilon)\alpha_p,$$
in which $\alpha_1 = k\sqrt{\frac{2}{\pi}}$ and $\alpha_2 = k$.

2. *The mapping $T : \mathbb{R}^n \to \mathbb{R}^k$ requires $O(n \log n + \min\{k\varepsilon^{-2} \log n, e^{p-4} \log^{p+1} k\})$ time to compute each matrix-vector multiplication.*

Chapter 3

Random projections for linear and integer programming

3.1 Restricted Linear Membership problems

Linear Programming (LP) is one of the most important and well-studied branches of optimization. An LP problem can be written in the following normal form

$$\max\{c^T x \mid Ax = b,\ x \geq 0\}.$$

It is well-known that it can be reduced (via an easy bisection argument) to LP feasibility problems, defined as follows:

> Linear Feasibility Problem(LFP). Given $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$. Decide whether there exists $x \in \mathbb{R}^n_+$ such that $Ax = b$.

We assume that m and n are very large integer numbers. Furthermore, as most other standard LPs, we also assume that A is a full row-rank matrix with $m \leq n$.

LFP problems can obviously be solved using the simplex method. Despite the fact that simplex methods are often very efficient in practice, there are instances for which the methods run in exponential time. On the other hand, polynomial time algorithms such as interior point methods are known to scale poorly, in practice, on several classes of instances. In any case, when m and n are huge, these methods fail to solve LPF. Our purpose is to use random projections to reduce considerably either m or n to the extend that traditional methods can apply.

Note that, if $a_1, \ldots, a_n$ are the column vectors of A, then the LPF is equivalent to finding $x \geq 0$ such that b is a non-negative linear combination of $a_1, \ldots, a_n$. In other words, the LPF is equivalent to the following cone membership problem:

> Cone Membership (CM). Given $b, a_1, \ldots, a_n \in \mathbb{R}^m$, decide whether $b \in \mathsf{cone}\{a_1, \ldots, a_n\}$.

It is known from the Johnson-Lindenstrauss lemma that there is a linear mapping $T : \mathbb{R}^m \to \mathbb{R}^k$, where $k \ll m$, such that the pairwise distances between all vector pairs (a_i, a_j) undergo low distortion. We are now stipulating that the complete distance graph is a reasonable representation of the intuitive notion of "shape". Under this hypothesis, it is reasonable to expect that the image of $C = \mathsf{cone}(a_1, \ldots, a_n)$ under T has approximately the same shape as C.

Thus, given an instance of CM, we expect to be able to "approximately solve" a much smaller (randomly projected) instance instead. Notice that since CM is a decision problem, "approximately" really refers to a randomized algorithm which is successful with high probability.

The LFP can be viewed as a special case of the *restricted linear membership problem*, which is defined as follows:

> Restricted Linear Membership (RLM). Given $b, a_1, \ldots, a_n \in \mathbb{R}^m$ and $X \subseteq \mathbb{R}^n$, decide whether $b \in \mathsf{lin}_X(a_1, \ldots, a_n)$, i.e. whether $\exists \lambda \in X$ s.t. $b = \sum_{i=1}^{n} \lambda_i a_i$.

RLM includes several very important classes of membership problems, such as

- When $X = \mathbb{R}^n_+$(or $\mathbb{R}^n_+ \cap \{\sum_{i=1}^n x_i = 1\}$), we have the cone membership problem (or convex hull membership problem), which corresponds to Linear Programming.
- When $X = \mathbb{Z}^n$ (or $\{0,1\}^n$), we have the integer (binary) cone membership problem (corresponding to Integer and Binary Linear Programming - ILP).
- When X is a convex set, we have the convex linear membership problem.
- When $n = d^2$ and X is the set of $d \times d-$positive semidefinite matrices, we have the semidefinite membership problem (corresponding to Semidefinite Programming- SDP).

Notation wise, every norm $\|\cdot\|$ is Euclidean unless otherwise specified, and we shall denote by A^{c} the complement of an event A. Moreover, we will implicitly assume (WLOG) that $a_1, \ldots, a_n, b, c$ are unit vectors.

The following lemma shows that the kernels of random projections are "concentrated around zero". This can be seen as a direct derivation from the definition of RP.

Corollary 3.1.1. *Let $T : \mathbb{R}^m \to \mathbb{R}^k$ be a random projection as in the definition 2.2.1 and let $x \in \mathbb{R}^m$ be a non-zero vector. Then we have*

$$\mathbb{P}(T(x) \neq 0) \geq 1 - 2e^{-\mathcal{C}k}. \tag{3.1}$$

for some constant $\mathcal{C} > 0$ (independent of m, k).

Proof. For any $\varepsilon \in (0, 1)$, we define the following events:

$$A = \big\{T(x) \neq 0\big\}$$
$$B = \big\{(1-\varepsilon)\|x\| \leq \|T(x)\| \leq (1+\varepsilon)\|x\|\big\}.$$

By Definition 2.2.1, it follows that $\mathbb{P}(B) \geq 1 - 2e^{-\mathcal{C}\varepsilon^2 k}$ for some constant $\mathcal{C} > 0$ independent of m, k, ε. On the other hand, $A^{\mathsf{c}} \cap B = \emptyset$, since otherwise, there is a mapping T_1 such that $T_1(x) = 0$ and $(1-\varepsilon)\|x\| \leq \|T_1(x)\|$, which altogether imply that $x = 0$ (a contradiction). Therefore, $B \subseteq A$, and we have $\mathbb{P}(A) \geq \mathbb{P}(B) \geq 1 - 2e^{-\mathcal{C}\varepsilon^2 k}$. This holds for all $0 < \varepsilon < 1$, so we have $\mathbb{P}(A) \geq 1 - 2e^{-\mathcal{C}k}$. □

Lemma 3.1.2. *Let $T : \mathbb{R}^m \to \mathbb{R}^k$ be a random projection as in the definition 2.2.1 and let $b, a_1, \ldots, a_n \in \mathbb{R}^m$. Then for any given vector $x \in \mathbb{R}^n$, we have:*

(i) If $b = \sum_{i=1}^n x_i a_i$ then $T(b) = \sum_{i=1}^n x_i T(a_i)$;

(ii) If $b \neq \sum_{i=1}^n x_i a_i$ then $\mathbb{P}\left[T(b) \neq \sum_{i=1}^n x_i T(a_i)\right] \geq 1 - 2e^{-\mathcal{C}k}$;

*(iii) If **for all** $y \in X \subseteq \mathbb{R}^n$ where $|X|$ is finite, we have $b \neq \sum_{i=1}^n y_i a_i$, then*

$$\mathbb{P}\left[\forall y \in X,\ T(b) \neq \sum_{i=1}^n y_i T(a_i)\right] \geq 1 - 2|X|e^{-\mathcal{C}k};$$

for some constant $\mathcal{C} > 0$ (independent of n, k).

Proof. Point (i) follows by linearity of T, and (ii) by applying Cor. 3.1.1 to $Ax - b$. For (iii), we have

$$\mathbb{P}\left[\forall y \in X,\ T(b) \neq \sum_{i=1}^n y_i T(a_i)\right] = \mathbb{P}\left[\bigcap_{y \in X} \big\{T(b) \neq \sum_{i=1}^n y_i T(a_i)\big\}\right]$$
$$= 1 - \mathbb{P}\left[\bigcup_{y \in X} \big\{T(b) \neq \sum_{i=1}^n y_i T(a_i)\big\}^{\mathsf{c}}\right] \geq 1 - \sum_{y \in X} \mathbb{P}\left[\big\{T(b) \neq \sum_{i=1}^n y_i T(a_i)\big\}^{\mathsf{c}}\right]$$
$$\text{[by (ii)]} \qquad \geq 1 - \sum_{y \in X} 2e^{-\mathcal{C}k} = 1 - 2|X|e^{-\mathcal{C}k},$$

as claimed. □

This lemma can be used to solve the RLM problem when the cardinality of the restricted set X is bounded by a polynomial in n. In particular, if $|X| < n^d$, where d is small w.r.t. n, then

$$\mathbb{P}\big\{T(b) \notin \mathsf{Lin}_X \{T(a_1), \ldots, T(a_n)\}\big\} \geq 1 - 2n^d e^{-\mathcal{C}k}. \tag{3.2}$$

Then by taking any k such that $k \geq \frac{1}{\mathcal{C}} \ln(\frac{2}{\delta}) + \frac{d}{\mathcal{C}} \ln n$, we obtain a probability of success of at least $1 - \delta$. We give an example to illustrate that such a bound for $|X|$ is natural in many different settings.

Example 3.1.3. *If $X = \{x \in \{0,1\}^n | \sum_{i=1}^n \alpha_i x_i \leq d\}$ for some d, where $0 < \alpha_i$ for all $1 \leq i \leq n$, then $|X| < n^{\overline{d}}$, where $\overline{d} = \max_{1 \leq i \leq n} \lfloor \frac{d}{\alpha_i} \rfloor$.*

To see this, let $\underline{\alpha} = \min_{1 \leq i \leq n} \alpha_i$; then $\sum_{i=1}^n x_i \leq \sum_{i=1}^n \frac{\alpha_i}{\underline{\alpha}} x_i \leq \frac{d}{\underline{\alpha}}$, which implies $\sum_{i=1}^n x_i \leq \overline{d}$. Therefore $|X| \leq \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{\overline{d}} < n^{\overline{d}}$, as claimed.

Lemma 3.1.2 also gives us an indication as to why estimating the probability that $T(b) \notin \mathsf{cone}\{T(a_1), \ldots, T(a_n)\}$ is not straightforward. This event can be written as an intersection of infinitely many sub-events $\{T(b) \neq \sum_{i=1}^n y_i T(a_i)\}$ where $y \in \mathbb{R}^n_+$; even if each of these occurs with high probability, their intersection might still be small. As these events are dependent, however, we still hope to find to find a useful estimation for this probability.

3.2 Projections of separating hyperplanes

In this section we show that if a hyperplane separates a point x from a closed and convex set C, then its image under a random projection T is also likely to separate $T(x)$ from $T(C)$. The separating hyperplane theorem applied to cones can be stated as follows.

Theorem 3.2.1 (Separating hyperplane theorem)**.** *Given $b \notin \mathsf{cone}\{a_1, \ldots, a_n\}$ where $b, a_1, \ldots, a_n \in \mathbb{R}^m$. Then there is $c \in \mathbb{R}^m$ such that $c^T b < 0$ and $c^T a_i \geq 0$ for all $i = 1, \ldots, n$.*

For simplicity, we will first work with *pointed cone*. Recall that a cone C is called pointed if and only if $C \cap -C = \{0\}$. The associated separating hyperplane theorem is obtained by replacing all $\geq$ inequalities by strict ones. Without loss of generality, we can assume that $\|c\| = 1$. From this theorem, it immediately follows that there is a positive ε_0 such that $c^T b < -\varepsilon_0$ and $c^T a_i > \varepsilon_0$ for all $1 \leq i \leq n$.

Proposition 3.2.2. *Given unit vectors $b, a_1, \ldots, a_n \in \mathbb{R}^m$ such that $b \notin \mathsf{cone}\{a_1, \ldots, a_n\}$. Let $\varepsilon > 0$ and $c \in \mathbb{R}^m$ with $\|c\| = 1$ be such that $c^T b < -\varepsilon$ and $c^T a_i \geq \varepsilon$ for all $1 \leq i \leq n$. Let $T : \mathbb{R}^m \to \mathbb{R}^k$ be a random projection as in the definition 2.2.1, then*

$$\mathbb{P}\big[T(b) \notin \mathsf{cone}\{T(a_1), \ldots, T(a_n)\}\big] \geq 1 - 4(n+1)e^{-\mathcal{C}\varepsilon^2 k}$$

for some constant $\mathcal{C}$ (independent of m, n, k, ε).

Proof. Let A be the event that both $(1-\varepsilon)\|c-x\|^2 \leq \|T(c-x)\|^2 \leq (1+\varepsilon)\|c-x\|^2$ and $(1-\varepsilon)\|c+x\|^2 \leq \|T(c+x)\|^2 \leq (1+\varepsilon)\|c+x\|^2$ hold for all $x \in \{b, a_1, \ldots, a_n\}$. By Definition 2.2.1, we have $\mathbb{P}(A) \geq 1 - 4(n+1)e^{-\mathcal{C}\varepsilon^2 k}$. For any random mapping T such that A occurs, we have

$$\begin{aligned}\langle T(c), T(b)\rangle &= \frac{1}{4}(\|T(c+b)\|^2 - \|T(c-b)\|^2) \\ &\leq \frac{1}{4}(\|c+b\|^2 - \|c-b\|^2) + \frac{\varepsilon}{4}(\|c+b\|^2 + \|c-b\|^2) \\ &= c^T b + \varepsilon < 0\end{aligned}$$

and similarly, for all $i = 1, \ldots, n$, we can derive $\langle T(c), T(a_i)\rangle \geq c^T a_i - \varepsilon \geq 0$. Therefore, by Thm. 3.2.1, $T(b) \notin \mathsf{cone}\{T(a_1), \ldots, T(a_n)\}$. □

From this proposition, it follows that the larger ε will provide us a better probability. The largest ε can be found by solving the following optimization problem.

Separating Coefficient Problem (SCP). Given $b \notin \mathsf{cone}\{a_1, \ldots, a_n\}$, find

$$\varepsilon = \max_{c,\varepsilon} \{\varepsilon |\, \varepsilon \geq 0, c^T b \leq -\varepsilon, c^T a_i \geq \varepsilon\}.$$

Note that ε can be extremely small when the cone C generated by $a_1, \ldots, a_n$ is almost non-pointed, i.e., the convex hull of $a_1, \ldots, a_n$ contains a point close to 0. Indeed, for any convex combination $x = \sum_i \lambda_i a_i$ with $\sum_i \lambda_i = 1$ of a_i's, we have:

$$\|x\| = \|x\| \cdot \|c\| \geq c^T x = \sum_{i=1}^{n} \lambda_i c^T a_i \geq \sum_{i=1}^{n} \lambda_i \varepsilon = \varepsilon.$$

Therefore, $\varepsilon \leq \min\{\|x\| \mid x \in \mathsf{conv}\{a_1, \ldots, a_n\}\}$.

3.3 Projection of minimum distance

In this section we show that if the distance between a point x and a closed set is positive, it remains positive with high probability after applying a random projection. First, we consider the following problem.

Convex Hull Membership (CHM). Given $b, a_1, \ldots, a_n \in \mathbb{R}^m$, decide whether $b \in \mathsf{conv}\{a_1, \ldots, a_n\}$.

Applying random projections, we obtain the following proposition:

Proposition 3.3.1. *Given $a_1, \ldots, a_n \in \mathbb{R}^m$, let $C = \mathsf{conv}\{a_1, \ldots, a_n\}$, $b \in \mathbb{R}^m$ such that $b \notin C$, $d = \min_{x \in C} \|b - x\|$ and $D = \max_{1 \le i \le n} \|b - a_i\|$. Let $T : \mathbb{R}^m \to \mathbb{R}^k$ be a random projection as in the definition 2.2.1. Then*

$$\mathbb{P}\big[T(b) \notin T(C)\big] \ge 1 - 2n^2 e^{-\mathcal{C}\varepsilon^2 k} \tag{3.3}$$

for some constant $\mathcal{C}$ (independent of m, n, k, d, D) and $\varepsilon < \frac{d^2}{D^2}$.

We will not prove this proposition. Instead we will prove the following generalized result concerning the separation of two convex hulls under random projections.

Proposition 3.3.2. *Given two disjoint polytopes $C = \mathsf{conv}\{a_1, \ldots, a_n\}$ and $C^* = \mathsf{conv}\{a_1^*, \ldots, a_p^*\}$ in $\mathbb{R}^m$, let $d = \min_{x \in C, y \in C^*} \|x - y\|$ and $D = \max_{1 \le i \le n, 1 \le j \le p} \|a_i - a_j^*\|$. Let $T : \mathbb{R}^m \to \mathbb{R}^k$ be a random projection. Then*

$$\mathbb{P}\big[T(C) \cap T(C^*) = \emptyset\big] \ge 1 - 2n^2 p^2 e^{-\mathcal{C}\varepsilon^2 k} \tag{3.4}$$

for some constant $\mathcal{C}$ (independent of m, n, p, k, d, D) and $\varepsilon < \frac{d^2}{D^2}$.

Proof. Let S_ε be the event that both $(1-\varepsilon)\|x-y\|^2 \le \|T(x-y)\|^2 \le (1+\varepsilon)\|x-y\|^2$ and $(1-\varepsilon)\|x+y\|^2 \le \|T(x+y)\|^2 \le (1+\varepsilon)\|x+y\|^2$ hold for all $x, y \in \{a_i - a_j^* \mid 1 \le i \le n, 1 \le j \le p\}$.

Assume S_ε occurs. Then for all reals $\lambda_i \ge 0$ with $\sum_{i=1}^n \lambda_i = 1$ and $\gamma_j \ge 0$ with $\sum_{j=1}^p \gamma_j = 1$, we have:

$$\begin{aligned}
& \|\sum_{i=1}^n \lambda_i T(a_i) - \sum_{j=1}^p \gamma_j T(a_j^*)\|^2 = \|\sum_{i=1}^n \sum_{j=1}^p \lambda_i \gamma_j T(a_i - a_j^*)\|^2 \\
= \;& \sum_{i=1}^n \sum_{j=1}^p \lambda_i^2 \gamma_j^2 \|T(a_i - a_j^*)\|^2 + 2 \sum_{(i,j) \ne (i',j')} \lambda_i \gamma_j \lambda_{i'} \gamma_{j'} \langle T(a_i - a_j^*), T(a_{i'} - a_{j'}^*) \rangle \\
= \;& \sum_{i=1}^n \sum_{j=1}^p \lambda_i^2 \gamma_j^2 \|T(a_i - a_j^*)\|^2 + \frac{1}{2} \sum_{(i,j) \ne (i',j')} \lambda_i \gamma_j \lambda_{i'} \gamma_{j'} \Big(\|T(a_i - a_j^* + a_{i'} - a_{j'}^*)\|^2 - \|T(a_i - a_j^* - a_{i'} + a_{j'}^*)\|^2 \Big) \\
\ge \;& (1-\varepsilon) \sum_{i=1}^n \sum_{j=1}^p \lambda_i^2 \gamma_j^2 \|a_i - a_j^*\|^2 + \\
& + \frac{1}{2} \sum_{(i,j) \ne (i',j')} \lambda_i \gamma_j \lambda_{i'} \gamma_{j'} \Big((1-\varepsilon) \big\|a_i - a_j^* + a_{i'} - a_{j'}^*\big\|^2 - (1+\varepsilon) \|a_i - a_j^* - a_{i'} + a_{j'}^*\|^2 \Big) \\
= \;& \|\sum_{i=1}^n \lambda_i a_i - \sum_{j=1}^p \gamma_j a_j^*\|^2 - \varepsilon \Big(\sum_{i=1}^n \sum_{j=1}^p \lambda_i^2 \gamma_j^2 \|a_i - a_j^*\|^2 + \sum_{(i,j) \ne (i',j')} \lambda_i \gamma_j \lambda_{i'} \gamma_{j'} \big(\|a_i - a_j^*\|^2 + \|a_{i'} - a_{j'}^*\|^2 \big) \Big).
\end{aligned}$$

From the definitions of d and D, we have:

$$\|\sum_{i=1}^n \lambda_i T(a_i) - \sum_{j=1}^p \gamma_j T(a_j^*)\|^2 \ge d^2 - \varepsilon D^2 \Big(\sum_{i=1}^n \sum_{j=1}^p \lambda_i^2 \gamma_j^2 + 2 \sum_{(i,j) \ne (i',j')} \lambda_i \gamma_j \lambda_{i'} \gamma_{j'} \Big) \ge d^2 - \varepsilon D^2 > 0$$

due to the choice of $\varepsilon < \frac{d^2}{D^2}$. In summary, if S_ε occurs, then $T(C)$ and $T(C^*)$ are disjoint. Thus, by the definition of random projection and the union bound,

$$\mathbb{P}(T(C) \cap T(C^*) = \emptyset) \geq \mathbb{P}(S_\varepsilon) \geq 1 - 2\big(np\big)^2 e^{-\mathcal{C}\varepsilon^2 k}$$

for some constant $\mathcal{C} > 0$. □

Now we assume that $b, c, a_1, \ldots, a_n$ are all unit vectors. In order to deal with the CM problem, we consider the so-called A-norm of $x \in \mathsf{cone}\{a_1, \ldots, a_n\}$ as $\|x\|_A = \min\big\{ \sum_{i=1}^n \lambda_i \;\big|\; \lambda \geq 0 \wedge x = \sum_{i=1}^n \lambda_i a_i \big\}$. For each $x \in \mathsf{cone}\{a_1, \ldots, a_n\}$, we say that $\lambda \in \mathbb{R}^n_+$ yields a *minimal A-representation* of x if and only if $\sum_{i=1}^n \lambda_i = \|x\|_A$. We define $\mu_A = \max\{\|x\|_A \mid x \in \mathsf{cone}\{a_1, \ldots, a_n\} \wedge \|x\| \leq 1\}$; then, for all $x \in \mathsf{cone}\{a_1, \ldots, a_n\}$, $\|x\| \leq \|x\|_A \leq \mu_A \|x\|$. In particular $\mu_A \geq 1$. Note that μ_A serves as a measure of worst-case distortion when we move from Euclidean to $\|\cdot\|_A$ norm.

Theorem 3.3.3. *Given unit vectors $b, a_1, \ldots, a_n \in \mathbb{R}^m$ such that $b \notin C = \mathsf{cone}\{a_1, \ldots, a_n\}$, let $d = \min_{x \in C} \|b - x\|$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ be a random projection as in the definition 2.2.1. Then:*

$$\mathbb{P}(\, T(b) \notin \mathsf{cone}\{T(a_1), \ldots, T(a_n)\}\,) \geq 1 - 2n(n+1)e^{-\mathcal{C}\varepsilon^2 k} \tag{3.5}$$

for some constant $\mathcal{C}$ (independent of m, n, k, d), in which $\varepsilon = \frac{d^2}{\mu_A^2 + 2\sqrt{1-d^2}\mu_A + 1}$.

Proof. For any $0 < \varepsilon < 1$, let S_ε be the event that both $(1-\varepsilon)\|x-y\|^2 \leq \|T(x-y)\|^2 \leq (1+\varepsilon)\|x-y\|^2$ and $(1-\varepsilon)\|x+y\|^2 \leq \|T(x+y)\|^2 \leq (1+\varepsilon)\|x+y\|^2$ hold for all $x, y \in \{b, a_1, \ldots, a_n\}$. By definition of random projection and the union bound, we have

$$\mathbb{P}(S_\varepsilon) \geq 1 - 4\binom{n+1}{2} e^{-\mathcal{C}\varepsilon^2 k} = 1 - 2n(n+1)e^{-\mathcal{C}\varepsilon^2 k}$$

for some constant $\mathcal{C}$ (independent of m, n, k, d). We will prove that if S_ε occurs, then we have $T(b) \notin \mathsf{cone}\{T(a_1), \ldots, T(a_n)\}$. Assume that S_ε occurs. Consider an arbitrary $x \in \mathsf{cone}\{a_1, \ldots, a_n\}$ and let $\sum_{i=1}^n \lambda_i a_i$ be the minimal A-representation of x. Then we have:

$$\begin{aligned}
& \|T(b) - T(x)\|^2 = \|T(b) - \sum_{i=1}^n \lambda_i T(a_i)\|^2 \\
= \;& \|T(b)\|^2 + \sum_{i=1}^n \lambda_i^2 \|T(a_i)\|^2 - 2\sum_{i=1}^n \lambda_i \langle T(b), T(a_i)\rangle + 2\sum_{1 \leq i < j \leq n} \lambda_i \lambda_j \langle T(a_i), T(a_j)\rangle \\
= \;& \|T(b)\|^2 + \sum_{i=1}^n \lambda_i^2 \|T(a_i)\|^2 + \sum_{i=1}^n \frac{\lambda_i}{2}(\|T(b-a_i)\|^2 - \|T(b+a_i)\|^2) + \sum_{1 \leq i < j \leq n} \frac{\lambda_i \lambda_j}{2}(\|T(a_i+a_j)\|^2 - \|T(a_i-a_j)\|^2) \\
\geq \;& (1-\varepsilon)\|b\|^2 + (1-\varepsilon)\sum_{i=1}^n \lambda_i^2 \|a_i\|^2 + \sum_{i=1}^n \frac{\lambda_i}{2}((1-\varepsilon)\|b-a_i\|^2 - (1+\varepsilon)\|b+a_i\|^2) \\
& \qquad + \sum_{1 \leq i < j \leq n} \frac{\lambda_i \lambda_j}{2}((1-\varepsilon)\|a_i+a_j\|^2 - (1+\varepsilon)\|a_i-a_j\|^2),
\end{aligned}$$

because of the assumption that S_ε occurs. Since $\|b\| = \|a_1\| = \cdots = \|a_n\| = 1$, the RHS can be written as

$$\begin{aligned}
&\|b - \sum_{i=1}^n \lambda_i a_i\|^2 - \varepsilon\Big(1 + \sum_{i=1}^n \lambda_i^2 + 2\sum_{i=1}^n \lambda_i + 2\sum_{i<j} \lambda_i\lambda_j\Big) \\
=&\|b - \sum_{i=1}^n \lambda_i a_i\|^2 - \varepsilon\big(1 + \sum_{i=1}^n \lambda_i\big)^2 \\
=&\|b - x\|^2 - \varepsilon\big(1 + \|x\|_A\big)^2
\end{aligned}$$

Denote by $\alpha = \|x\|$ and let p be the projection of b to $\mathsf{cone}\{a_1, \ldots, a_n\}$, which implies $\|b - p\| = \min\{\|b - x\| \mid x \in \mathsf{cone}\{a_1, \ldots, a_n\}\}$.

Claim. For all b, x, α, p given above, we have $\|b - x\|^2 \geq \alpha^2 - 2\alpha\|p\| + 1$.

By this claim (proved later), we have:

$$\begin{aligned}
\|T(b) - T(x)\|^2 &\geq \alpha^2 - 2\alpha\|p\| + 1 - \varepsilon\big(1 + \|x\|_A\big)^2 \\
\geq \alpha^2 - 2\alpha\|p\| + 1 - \varepsilon\big(1 + \mu_A\alpha\big)^2 &= \big(1 - \varepsilon\mu_A^2\big)\alpha^2 - 2\big(\|p\| + \varepsilon\mu_A\big)\alpha + (1 - \varepsilon).
\end{aligned}$$

The last expression can be viewed as a quadratic function with respect to α. We will prove this function is nonnegative for all $\alpha \in \mathbb{R}$. This is equivalent to

$$\begin{aligned}
&\big(\|p\| + \varepsilon\mu_A\big)^2 - \big(1 - \varepsilon\mu_A^2\big)(1 - \varepsilon) \leq 0 \\
\Leftrightarrow\ &\big(\mu_A^2 + 2\|p\|\mu_A + 1\big)\varepsilon \leq 1 - \|p\|^2 \\
\Leftrightarrow\ &\varepsilon \leq \frac{1 - \|p\|^2}{\mu_A^2 + 2\|p\|\mu_A + 1} = \frac{d^2}{\mu_A^2 + 2\|p\|\mu_A + 1},
\end{aligned}$$

which holds for the choice of ε as in the hypothesis. In summary, if the event S_ε occurs, then $\|T(b) - T(x)\|^2 > 0$ for all $x \in \mathsf{cone}\{a_1, \ldots, a_n\}$, i.e. $T(x) \notin \mathsf{cone}\{T(a_1), \ldots, T(a_n)\}$. Thus,

$$\mathbb{P}(T(b) \notin TC) \geq \mathbb{P}(S_\varepsilon) \geq 1 - 2n(n+1)e^{-\mathcal{C}\varepsilon^2 k}$$

as claimed. □

Proof of the claim. If $x = 0$ then the claim is trivially true, since $\|b - x\|^2 = \|b\|^2 = 1 = \alpha^2 - 2\alpha\|p\| + 1$. Hence we assume $x \neq 0$. First consider the case $p \neq 0$. By Pythagoras' theorem, we must have $d^2 = 1 - \|p\|^2$. We denote by $z = \frac{\|p\|}{\alpha}x$, then $\|z\| = \|p\|$. Set $\delta = \frac{\alpha}{\|p\|}$,

we have

$$\begin{aligned}\|b-x\|^2 &= \|b-\delta z\|^2\\ &= (1-\delta)\|b\|^2 + (\delta^2-\delta)\|z\|^2 + \delta\|b-z\|^2\\ &= (1-\delta) + (\delta^2-\delta)\|p\|^2 + \delta\|b-z\|^2\\ &\geq (1-\delta) + (\delta^2-\delta)\|p\|^2 + \delta d^2\\ &= (1-\delta) + (\delta^2-\delta)\|p\|^2 + \delta(1-\|p\|^2)\\ &= \delta^2\|p\|^2 - 2\delta\|p\|^2 + 1 = \alpha^2 - 2\alpha\|p\| + 1.\end{aligned}$$

Next, we consider the case $p = 0$. In this case we have $b^T(x) \leq 0$ for all $x \in \mathsf{cone}\{a_1, \ldots, a_n\}$. Indeed, for an arbitrary $\delta > 0$,

$$0 \leq \frac{1}{\delta}(\|b-\delta x\|^2 - 1) = \frac{1}{\delta}(1 + \delta^2\|x\|^2 - 2\delta b^T x - 1) = \delta\|x\|^2 - 2b^T x$$

which tends to $-2b^T x$ when $\delta \to 0^+$. Therefore

$$\|b-x\|^2 = 1 - 2b^T x + \|x\|^2 \geq \|x\|^2 + 1 = \alpha^2 - 2\alpha\|p\| + 1,$$

which proves the claim. □

3.4 Certificates for projected problems

In this section, we will prove that, under certain assumptions, the solutions we obtain by solving the projected feasibility problem are not likely to be solutions for the original problems. They are unfortunately negative results, which reflects the fact that we cannot simply solve the projected problem to obtain a solution but need a smarter way to deal with it.

In the first proposition, we assume that a solution is found uniformly in the feasible set of the projected problem. However, this assumption does not hold if we use some popular methods (such as simplex) because in these cases we rather end up with extreme points. We make use of this observation in our second proposition.

Recall that, we want to study the relationship between the linear feasibility problem (LFP):

> Decide whether there exists $x \in \mathbb{R}^n$ such that $Ax = b \wedge x \geq 0$.

and it projected version (PLFP):

> Decide whether there exists $x \in \mathbb{R}^n$ such that $TAx = Tb \wedge x \geq 0$,

where $T \in \mathbb{R}^{k\times m}$ is a real matrix. To keep all the following discussions meaningful, we assume that $k < m$ and the LFP is feasible. Here we use the terminology *certificate* to indicate a solution x that verifies the feasibility of the associated problem.

Proposition 3.4.1. *Assume that b belongs to the interior of $\mathsf{cone}(A)$ and x^* is uniformly chosen from the feasible set of the PLFP. Then x^* is almost surely not a certificate for the LFP.*

In a formal way, let $\mathcal{O} = \{x \geq 0 \mid Ax = b\}$ and $\mathcal{P} = \{x \geq 0 \mid TAx = Tb\}$ denote the feasible sets for the original and projected problems, and let x^* be uniformly distributed on $\mathcal{P}$, i.e. $\mathcal{P}$ is equipped with a uniform probability measure μ. For each $v \in \mathsf{ker}(T)$, let

$$\mathcal{O}_v = \{x \geq 0 \mid Ax - b = v\} \cap \mathcal{P}$$

Notice that here $\mathcal{O}_0 = \mathcal{O}$. We need to prove that $\mathrm{Prob}(x^* \in \mathcal{O}) = 0$.

Proof. Assume for contradiction that

$$\mathrm{Prob}(x^* \in \mathcal{O}) = p > 0.$$

We will prove that there exists $\delta > 0$ and a family $\mathcal{V}$ of infinitely many $v \in \mathsf{ker}(T)$ such that $\mathrm{Prob}(x^* \in \mathcal{O}_v) \geq \delta > 0$. Since $(\mathcal{O}_v)_{v\in\mathcal{V}}$ is a family of disjoint sets, we deduce that $\mathrm{Prob}\left(x^* \in \bigcup\limits_{v\in\mathcal{V}} \mathcal{O}_v\right) \geq \sum\limits_{v\in\mathcal{V}} \delta = +\infty$, leading to a contradiction.

Because $\mathsf{dim}(\mathsf{ker}(T)) \geq 1$, then $\mathsf{ker}(T)$ contains a segment $[-u, u]$. Furthermore, since $0 \in \mathrm{int}\{Ax - b \mid x \geq 0\}$ (due to the assumption that b belongs to the interior of $\mathsf{cone}(A)$), we can choose $\|u\|$ small enough such that $[-u, u]$ also belongs to $\{Ax - b \mid x \geq 0\}$.

Also due to the assumption that b belongs to the interior of $\mathsf{cone}(A)$, there exists $\bar{x} > 0$ such that $A\bar{x} = b$. Let $\hat{x} \in \mathbb{R}^n$ be such that $A\hat{x} = -u$. There always exists an $\bar{N} > 0$ large enough such that $2\hat{x} \leq \bar{N}\bar{x}$ (since $\bar{x} > 0$).

For all $N \geq \bar{N}$ and for all $x \in \mathcal{O}$, we denote $x'_N = \frac{\bar{x}+x}{2} - \frac{1}{N}\hat{x}$. Then we have $Ax'_N = b - \frac{1}{N}A\hat{x} = b + \frac{u}{N}$ and $x'_N = \frac{x}{2} + (\frac{\bar{x}}{2} - \frac{\hat{x}}{N}) \geq 0$. Therefore,

$$\frac{\bar{x} + \mathcal{O}}{2} - \frac{1}{N}\hat{x} \subseteq \mathcal{O}_{\frac{u}{N}}$$

which implies that, for all $N \geq \bar{N}$,

$$\mathrm{Prob}(x^* \in \mathcal{O}_{\frac{u}{N}}) = \mu(\mathcal{O}_{\frac{u}{N}}) \geq \mu(\frac{\bar{x} + \mathcal{O}}{2}) \geq c\mu(\mathcal{O}) = cp > 0$$

for some constant $c > 0$. The proposition is proved. □

Proposition 3.4.2. *Assume that b does not belong to the boundary of $\mathsf{cone}(A_B)$ for any basis A_B of A and x^* is an extreme point of the projected feasible set. Then x^* is not a certificate for the LFP.*

For consistency, we use the same notations $\mathcal{O}, \mathcal{P}$ as before to denote the feasible sets of the LFP and its projected problem. We also denote by $\mathcal{O}^*, \mathcal{P}^*$ their vertex sets, respectively.

Proof. For contradiction, we assume that x^* is also a certificate for the LFP. Then we claim that x^* is an extreme point of the LFP feasible set, i.e. $x^* \in \mathcal{O}^*$. Indeed, if this does not hold, then there are two distinct $x^1, x^2 \in \mathcal{O}$ such that $x^* = \frac{1}{2}(x^1 + x^2)$. However, since $\mathcal{O} \subseteq \mathcal{P}$, then both x^1, x^2 belong to $\mathcal{P}$, which contradicts to the assumption that x^* is an extreme point of the projected feasible set.

For that reason, we can write $x^* = (x^*_B, x^*_N)$, where $x^*_B = (A_B)^{-1}b$ is a basis solution and $x^*_N = 0$ is a non-basis solution. It then follows that $b \in \mathsf{cone}(A_B)$, and due to our first assumption, $b \in \text{int}(\mathcal{C}_B)$. Let $b = A_B\overline{x}$ for some $\overline{x} > 0$. Since $A_B\overline{x} = A_B x^*_B$, it turns out that $x^*_B = \overline{x} > 0$ (due to the non-singularity of A_B). Now we have a contradiction, since every extreme point of the projected LFP has at most k non-zero components, but x^* has exactly m non-zero components ($m > k$). The proof is finished. □

Please notice that the assumptions in the above two propositions hold almost surely when the instances A, b are uniformly generated. This explains the results we obtained for random instances.

Now we consider the Integer Feasibility Problem (IFP)

DECIDE WHETHER THERE EXISTS $x \in \mathbb{Z}^n_+$ SUCH THAT $Ax = b$

and it projected version (PIFP):

DECIDE WHETHER THERE EXISTS $x \in \mathbb{Z}^n$ SUCH THAT $TAx = Tb \wedge x \geq 0$,

where $T \in \mathbb{R}^{k\times m}$ is a real matrix. We will prove that

Proposition 3.4.3. *Assume that T is sampled from a probability distribution with bounded Lebesgue density, and the IFP is feasible. Then any certificate for the projected IFP will almost surely be a certificate for the original IFP.*

We first prove the following simple lemma:

Lemma 3.4.4. *Let ν be a probability distribution on $\mathbb{R}^{mk}$ with bounded Lebesgue density. Let $Y \subseteq \mathbb{R}^m$ be an at most countable set such that $0 \notin Y$. Then, for a random projection $T : \mathbb{R}^m \to \mathbb{R}^k$ sampled from ν, we have $0 \notin T(Y)$ almost surely, i.e. $\mathbb{P}\big(0 \notin T(Y)\big) = 1$.*

Proof. Let f be the Lebesgue density of ν. For any $0 \neq y \in Y$, consider the set $\mathcal{E}_y = \{T : \mathbb{R}^m \to \mathbb{R}^k \mid T(y) = 0\}$. If we regard each $T : \mathbb{R}^m \to \mathbb{R}^1$ as a vector $t \in \mathbb{R}^{mk}$, then $\mathcal{E}_y$ is a hyperplane $\{t \in \mathbb{R}^{mk} | \; y \cdot t = 0\}$ and we have

$$\mathbb{P}(T(y) = 0) = \nu(\mathcal{E}_y) = \int_{\mathcal{E}_y} f d\mu \leq \|f\|_\infty \int_{\mathcal{E}_y} d\mu = 0$$

where μ denotes the Lebesgue measure on $\mathbb{R}^m$. The proof then follows by the countability of Y. □

Proof of the Proposition. Assume that x^* is a (integer) certificate for the projected IFP. Let $y^* = Ax^* - b$ and let $Z = \{Ax - b \mid x \in \mathbb{Z}^n_+\}$. Then 0 belongs to Z due to the feasibility of the original IFP. Moreover, Z is countable, so the above lemma implies that (applied on $Y = Z \setminus \{0\}$)

$$\mathsf{ker}(T) \cap Z = \{0\} \quad \text{almost surely.}$$

However, y^* belongs to both $\mathsf{ker}(T)$ and Z, therefore, $y^* = 0$ almost surely. In other words, x^* is a certificate for the IFP almost surely. □

3.5 Preserving Optimality in LP

Until now, we have only discussed about Linear Programming in the feasible form. In this section, we will directly consider the following LP:

$$(P) \quad \min\{c^T x | \; Ax = b, x \geq 0\},$$

in which A is an $\mathbb{R}^{m \times n}$ matrix ($m < n$) with full row rank. Its projected problems is given by

$$(P_T) \quad \min\{c^T x | \; TAx = Tb, x \geq 0\}.$$

Let $v(P)$ and $v(P_T)$ be the optimal objective value of the two problems (P) and (P_T), respectively. In this section we will show that $v(P) \approx v(P_T)$ with high probability. Our proof assumes that the feasible region of P is non-empty and bounded. Specifically, we assume that a constant $\theta > 0$ is given such that there exists an optimal solution x^* of P satisfying

$$\sum_{i=1}^{n} x_i^* < \theta. \tag{3.6}$$

For the sake of simplicity, we assume further that $\theta \geq 1$. This assumption is used to control the excessive flatness of the involved cones, which is required in the projected separation argument.

3.5.1 Transforming the cone membership problems

In this subsection, we will explain the idea of transforming a cone into another cone, so that the cone membership problem becomes easier to solve by random projection.

Given a polyhedral cone

$$\mathcal{K} = \left\{ \sum_{i=1}^{n} C_i x_i \;\middle|\; x \in \mathbb{R}^n_+ \right\}$$

in which $C_1, \ldots, C_n$ are column vectors of an $m \times n$ matrix C. For any $u \in \mathbb{R}^m$, we consider the following transformation ϕ_u, defined by:

$$\phi_u(\mathcal{K}) := \left\{ \sum_{i=1}^{n} \left(C_i - \frac{1}{\theta} u \right) x_i \;\middle|\; x \in \mathbb{R}^n_+ \right\}.$$

In particular, ϕ_u moves the origin in the direction u by a step $1/\theta$. For θ defined in the equation (3.6), we also consider the following set

$$\mathcal{K}_\theta = \left\{ \sum_{i=1}^{n} C_i x_i \;\middle|\; x \in \mathbb{R}^n_+ \wedge \sum_{i=1}^{n} x_i < \theta \right\}.$$

K_θ can be seen as a set truncated from K. We shall show that ϕ_u preserves the membership of u in the "truncated cone" $\mathcal{K}_\theta$. Moreover, ϕ_u, when applied to K_θ, will result in a more acute cone, which is easier for us to work with.

Lemma 3.5.1. *For any $u \in \mathbb{R}^m$, we have that $u \in \mathcal{K}_\theta$ if and only if $u \in \phi_u(\mathcal{K})$.*

Proof. For all $1 \leq i \leq n$, let $C'_i = C_i - \frac{1}{\theta} u$.

($\Rightarrow$) If $u \in \mathcal{K}_\theta$, then there exists $x \in \mathbb{R}^n_+$ such that $u = \sum_{i=1}^{n} C_i x_i$ and $\sum_{i=1}^{n} x_i < \theta$. Then $u \in \phi_u(\mathcal{K})$ because it can be written as $\sum_{i=1}^{n} C'_i x'_i$ with $x' = \frac{x}{1 - \frac{1}{\theta} \sum_{j=1}^{n} x_j}$. Note that, due to the assumption that $\sum_{j=1}^{n} x_j < \theta$, $x' \geq 0$ indeed.

($\Leftarrow$) If $u \in \phi_u(\mathcal{K})$, then there exists $x \in \mathbb{R}^n_+$ such that $u = \sum_{i=1}^{n} C'_i x_i$. Then u can also be

written as $\sum_{i=1}^{n} C'_i x'_i$, where $x'_i = \frac{x_i}{1+\frac{1}{\theta}\sum_{j=1}^{n} x_j}$. Note that, $\sum_{i=1}^{n} x'_i < \theta$ because

$$\sum_{i=1}^{n} x'_i = \frac{\sum_{i=1}^{n} x_i}{1 + \frac{1}{\theta}\sum_{i=1}^{n} x_i} < \theta,$$

which implies that $u \in K_\theta$.

□

Note that this result is still valid when the transformation ϕ_u is only applied to a subset of columns of C. Given an index set $I \subseteq \{1, \ldots, n\}$, we define $\forall i \leq n$:

$$C_i^I = \begin{cases} C_i - \frac{1}{\theta}u & \text{if } i \in I \\ C_i & \text{otherwise.} \end{cases}$$

We extend ϕ_u to

$$\phi_u^I(\mathcal{K}) = \left\{ \sum_{i=1}^{n} C_i^I x_i \;\middle|\; x \in \mathbb{R}_+^n \right\}, \tag{3.7}$$

and define

$$\mathcal{K}_\theta^I = \left\{ \sum_{i=1}^{n} C_i x_i \;\middle|\; x \in \mathbb{R}_+^n \wedge \sum_{i \in I} x_i < \theta \right\}.$$

The following corollary is proved in the same way as Lemma 3.5.1, in which ϕ_u is replaced by ϕ_u^I.

Corollary 3.5.2. *For any $u \in \mathbb{R}^m$ and $I \subseteq \{1, \ldots, n\}$, we have $u \in \mathcal{K}_\theta^I$ if and only if $u \in \phi_u^I(\mathcal{K})$.*

3.5.2 The main approximate theorem

Given an LFP instance $Ax = b \wedge x \geq 0$, where A is an $m \times n$ matrix and T is a $k \times m$ random projector. From the previous section, we know that

$$\exists x \geq 0\ (Ax = b) \quad \Leftrightarrow \quad \exists x \geq 0\ (TAx = Tb)$$

with high probability. We remark that this also holds for a $(k+h) \times m$ random projector of the form

$$\begin{pmatrix} I_h & 0 \\ & T \end{pmatrix},$$

where T is a $k \times m$ random projection. This allows us to claim the feasibility equivalence w.o.p. even when we only want to project a subset of rows of A. In the following, we will use this observation to handle constraints and objective function separately. In particular, we only project the constraints while keeping objective function unchanged.

If we add the constraint $\sum_{i=1}^{n} x_i \leq \theta$ to the problem P_T, we obtain the following:

$$P_{T,\theta} \equiv \min \left\{ c^\top x \;\middle|\; TAx = Tb \wedge \sum_{i=1}^{n} x_i \leq \theta \wedge x \in \mathbb{R}^n_+ \right\}. \tag{3.8}$$

The following theorem asserts that the optimal objective value of P can be well-approximated by that of $P_{T,\theta}$.

Theorem 3.5.3. *Assume $\mathcal{F}(P)$ is bounded and non-empty. Let y^* be an optimal dual solution of P of minimal Euclidean norm. Given $\delta > 0$, we have*

$$v(P) - \delta \leq v(P_{T,\theta}) \leq v(P), \tag{3.9}$$

with probability at least $1 - 4ne^{-\mathcal{C}\varepsilon^2 k}$, where $\varepsilon < \frac{\delta}{2(\theta+1)\eta}$, and η is $O(\|y^*\|)$.

Proof. First, we will briefly explain the idea of the proof. Since $v(P)$ is the optimal objective value of problem P, for any positive δ, the value $v(P) - \delta$ can not be attained. It means the following problem

$$Ax = b \wedge x \geq 0 \wedge cx \leq v(P) - \delta.$$

is infeasible. This problem can now be projected in such a way that it remains infeasible w.o.p. By writing this original problem as

$$\begin{pmatrix} c & 1 \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ s \end{pmatrix} = \begin{pmatrix} v(P) - \delta \\ b \end{pmatrix}, \text{ where } \begin{pmatrix} x \\ s \end{pmatrix} \geq 0, \tag{3.10}$$

and applying a random projection of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix}, \text{ where } T \text{ is a } k \times m \text{ random projection,}$$

we will obtain the following problem, which is supposed to be infeasible w.o.p.

$$\left.\begin{array}{rcl} cx + s & = & v(P) - \delta \\ TAx & = & Tb \\ x & \geq & 0 \end{array}\right\}. \tag{3.11}$$

The main idea is that, the prior information about the optimal solution x^* (i.e. $\sum_{i=1}^{n} x_i^* \leq \theta$), can now be added into this new problem. It does not change the feasibility of this problem,

but later can be used to transform the corresponding cone into a better one (more acute). Therefore, w.o.p., the problem

$$\left.\begin{array}{rcl} cx & \leq & v(P) - \delta \\ TAx & = & Tb \\ \sum_{i=1}^{n} x_i & \leq & \theta \\ x & \geq & 0 \end{array}\right\} \tag{3.12}$$

is infeasible. Hence we deduce that $cx \geq v(P) - \delta$ holds w.o.p. for any feasible solution x of the problem $P_{T,\theta}$, and that proves the LHS of Eq. (3.9). For the RHS, the proof is trivial since P_T is a relaxation of P with the same objective function.

Let

$$\tilde{A} = \begin{pmatrix} c^T & 1 \\ A & 0 \end{pmatrix}, \tilde{x} = \begin{pmatrix} x \\ s \end{pmatrix} \text{ and } \tilde{b} = \begin{pmatrix} v(P) - \delta \\ b \end{pmatrix}$$

Furthermore, let $\tilde{T} = \begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix}$. In the rest of the proof, we prove that $\tilde{b} \notin \mathsf{cone}(\tilde{A})$ iff $T\tilde{b} \notin \mathsf{cone}(T\tilde{A})$ w.o.p.

Let I be the set of indices of the first n columns of $\tilde{A}$. Consider the transformation $\phi_{\tilde{b}}^{I}$ as defined above, using a step $\frac{1}{\theta'}$ instead of $\frac{1}{\theta}$, in which $\theta' \in (\theta, \theta + 1)$. We define the following matrix:

$$A' = \begin{pmatrix} \tilde{A}_1 - \frac{1}{\theta'}\tilde{b} & \cdots & \tilde{A}_n - \frac{1}{\theta'}\tilde{b} & \tilde{A}_{n+1} \end{pmatrix}$$

Since Eq. (3.10) is infeasible, it is easy to verify that the system:

$$\left.\begin{array}{rcl} \tilde{A}\tilde{x} & = & \tilde{b} \\ \sum_{i=1}^{n} x_i & < & \theta' \\ \tilde{x} & \geq & 0 \end{array}\right\} \tag{3.13}$$

is also infeasible. Then, by Cor. 3.5.2, it follows that $\tilde{b} \notin \mathsf{cone}(A')$.

Let $y^* \in \mathbb{R}^m$ be an optimal dual solution of P of minimal Euclidean norm. By the strong duality theorem, we have $y^* A \leq c$ and $y^* b = v(P)$. We define $\tilde{y} = (1, -y^*)^\top$. We will prove that $\tilde{y} A' > 0$ and $\tilde{y}\tilde{b} < 0$. Indeed, since $\tilde{y}\tilde{A} = \begin{pmatrix} c - y^* A \\ 1 \end{pmatrix} \geq 0$ and $\tilde{y}\tilde{b} = v(P) - \delta - v(P) = -\delta < 0$, then we have

$$\tilde{y} A' = \begin{pmatrix} c - y^* A + \frac{\delta}{\theta'} \\ 1 \end{pmatrix} \geq \frac{\delta}{\theta'} \geq \frac{\delta}{\theta + 1} \text{ and } \tilde{y}\tilde{b} = -\delta, \tag{3.14}$$

which proves the claim.

Now we apply the scalar product preservation property and the union bound, to get

$$\|(\tilde{T}\tilde{y})\,(\tilde{T}A') - \tilde{y}\,A'\|_\infty \leq \varepsilon\eta \tag{3.15}$$

holds with probability at least $p = 1 - 4ne^{-\mathcal{C}\varepsilon^2 k}$. Here, η is the normalized constant

$$\eta = \max\left\{\|\tilde{y}\|.\|\tilde{b}\|,\ \max_{1\leq i\leq n}\ \|\tilde{y}\|.\|A'_i\|\right\},$$

in which $\eta = O(\theta\|y^*\|)$ (proof is given at the end). Let us now fix $\varepsilon = \frac{\delta}{2(\theta+1)\eta}$. By (3.14) and (3.15), we have, with probability at least p, the system

$$\begin{array}{rcl}(\tilde{T}\tilde{y})\,(\tilde{T}A') & \geq & 0\\ (\tilde{T}\tilde{y})\,(\tilde{T}\tilde{b}) & < & 0\end{array}$$

holds, which implies that the problem

$$\begin{array}{rcl}\tilde{T}A'\tilde{x} & = & \tilde{T}\tilde{b}\\ \tilde{x} & \geq & 0\end{array}$$

is infeasible (by Farkas' Lemma). By definition, $\tilde{T}A'\tilde{x} = \tilde{T}\tilde{A}\tilde{x} - \frac{1}{\theta'}\sum_{i=1}^{n} x_i\tilde{T}\tilde{b}$, which implies that

$$\left.\begin{array}{rcl}\tilde{T}\tilde{A}\tilde{x} & = & \tilde{T}\tilde{b}\\ \sum_{i=1}^{n}\tilde{x}_i & < & \theta'\\ \tilde{x} & \geq & 0\end{array}\right\}$$

is also infeasible with probability at least p (the proof is similar to that of Corollary 3.5.2). Therefore, with probability at least p, the following optimization problem:

$$\inf\left\{c^\top x \;\middle|\; TAx = Tb \wedge \sum_{i=1}^{n} x_i < \theta' \wedge x \in \mathbb{R}^n_+\right\}.$$

has its optimal value greater than $v(P) - \delta$. Since $\theta' > \theta$, it follows that with probability at least p, $v(P_{T,\theta}) \geq v(P) - \delta$, as claimed.

Proof of the claim $\eta = O(\theta\|y^*\|)$: We have

$$\begin{aligned}\|\tilde{b}\|^2 &= \|b\|^2 + (v(P) - \delta)^2\\ &\leq \|b\|^2 + 2v(P)^2\\ &= 1 + 2|c^\top x^*|\\ &\leq 1 + 2\|c\|_\infty.\|x^*\|_1 \quad \text{(by Hölder inequality)}\\ &\leq 1 + 2\theta \quad \text{(since } \|c\|_2 = 1 \text{ and } \textstyle\sum x_i^* \leq \theta\text{)}\\ &\leq 3\theta \quad \text{(by the assumption that } \theta \geq 1\text{)}.\end{aligned}$$

Therefore, we conclude that

$$\eta = \max\left\{\|\tilde{y}\|.\|\tilde{b}\|,\ \max_{1\leq i\leq n}\ \|\tilde{y}\|.\|A'_i\|\right\} = O(\theta\,\|y^*\|)$$

□

3.6 Computational results

Let T be the random projector, A the constraint matrix, b the RHS vector, and X either $\mathbb{R}^n_+$ in the case of LP and $\mathbb{Z}^n_+$ in the case of IP. We solve $Ax = b \wedge x \in X$ and $T(A)x = T(b) \wedge x \in X$ to compare accuracy and performance. In these results, A is dense. We generate (A, b) componentwise from three distributions: uniform on $[0, 1]$, exponential, gamma. For T, we only test the best-known type of projector matrix $T(y) = Py$, namely P is a random $k \times m$ matrix each component of which is independently drawn from a normal $\mathcal{N}(0, \frac{1}{\sqrt{k}})$ distribution. All problems were solved using CPLEX 12.6 on an Intel i7 2.70GHz CPU with 16.0 GB RAM. All the computational experiments were carried out in JuMP (a modeling language for Mathematical Programming in JULIA).

Because accuracy is guaranteed for feasible instances by Lemma 3.1.2 (i), we only test *infeasible* LP and IP feasibility instances. For every given size $m \times n$ of the constraint matrix, we generate 10 different instances, each of which is projected using 100 randomly generated projectors P. For each size, we compute the percentage of success, defined as an infeasible original problem being reduced to an infeasible projected problem. Performance is evaluated by recording the average user CPU time taken by CPLEX to solve the original problem, and, comparatively, the time taken to perform the matrix multiplication PA plus the time taken by CPLEX to solve the projected problem.

In the above computational results, we only report the actual solver execution time (in the case of the original problem) and matrix multiplication plus solver execution (in the case of the projected problem). Lastly, although Tables 3.1 tell a successful story, we obtained less satisfactory results with other distributions. Sparse instances yielded accurate but poorly performant results. So far, this seems to be a good practical method for dense LP/IP.

Table 3.1: LP: above, IP: below. *Acc.*: accuracy (% feas./infas. agreement), *Orig.*: original (CPU), *Proj.*: projected instances (CPU).

		Uniform			Exponential			Gamma		
m	n	Acc.	Orig.	Proj.	Acc.	Orig.	Proj.	Acc.	Orig.	Proj.
600	1000	99.5%	1.57s	0.12s	93.7%	1.66s	0.12s	94.6%	1.64s	0.11s
700	1000	99.5%	2.39s	0.12s	92.8%	2.19s	0.12s	93.1%	2.15s	0.11s
800	1000	99.5%	2.55s	0.11s	95.0%	2.91s	0.11s	97.3%	2.78 s	0.11s
900	1000	99.6%	3.49s	0.12s	96.1%	3.65s	0.13s	97.0%	3.57s	0.13s
1000	1500	99.5%	16.54s	0.20s	93.0%	18.10s	0.20s	91.2%	17.58s	0.20s
1200	1500	99.6%	22.46s	0.23s	95.7%	22.46s	0.20s	95.7%	22.58s	0.22s
1400	1500	100%	31.08s	0.24s	93.2%	35.24s	0.26s	95.0%	31.06s	0.23s
1500	2000	99.4%	48.89s	0.30s	91.3%	51.23s	0.31s	90.1%	51.08s	0.31
1600	2000	99.8%	58.36s	0.34s	93.8%	58.87s	0.34s	95.9%	60.35s	0.358s
500	800	100%	20.32s	4.15s	100%	4.69s	10.56s	100%	4.25s	8.11s
600	800	100%	26.41s	4.22s	100%	6.08s	10.45s	100%	5.96s	8.27s
700	800	100%	38.68s	4.19s	100%	8.25s	10.67s	100%	7.93s	10.28s
600	1000	100%	51.20s	7.84s	100%	10.31s	8.47s	100%	8.78s	6.90s
700	1000	100%	73.73s	7.86s	100%	12.56s	10.91s	100%	9.29s	8.43s
800	1000	100%	117.8s	8.74s	100%	14.11s	10.71s	100%	12.29s	7.58s
900	1000	100%	130.1s	7.50s	100%	15.58s	10.75s	100%	15.06s	7.65s
1000	1500	100%	275.8s	8.84s	100%	38.57s	22.62s	100%	35.70s	8.74s

Chapter 4

Random projections for convex optimization with linear constraints

4.1 Introduction

For motivation, we first consider the following convex optimization problem with linear constraints:

$$\min\{f(x) \mid Ax = b,\ x \in \mathcal{S}\},$$

in which $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, and $\mathcal{S} \subseteq \mathbb{R}^n$ is a convex set.

One example is from constrained model fitting, in which $Ax = b$ expresses the interpolation constraints, $x \in \mathcal{S}$ restricts the choices for model parameters to a certain domain and $f(x)$ indicates some cost (such as squared error $\| \ . \ \|^2$). Using bisection method, we can write this problem in the feasibility form as follows: Given $c \in \mathbb{R}$, decide whether the set $\{x|\ x \in \mathcal{S},\ f(x) \leq c_k,\ Ax = b\}$ is empty or not. Let denote by $\mathcal{C} = \{x \in \mathcal{S} \mid f(x) \leq c\}$. Since $f(x)$ is convex, $\mathcal{C}$ is also a convex set. Then, similar as the previous chapter, this feasibility problem can be viewed as a Convex Restricted Linear Membership problem (CRLM). Formally, we ask

> Convex RLM (CRLM). Given a closed convex set $\mathcal{C} \subseteq \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, decide whether the set
>
> $$\{x \in \mathcal{C} \mid Ax = b\} \quad \text{is empty or not.}$$

Instead of solving this problem, we can apply a random projection $T \in \mathbb{R}^{d\times m}$ to its linear constraints and study the projected version:

PROJECTED CRLM (PCRLM). Given a convex set $\mathcal{C} \subseteq \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, decide whether the set

$$\{x \in \mathcal{C} \mid TAx = Tb\} \quad \text{is empty or not.}$$

Usually, d is selected so that it is much smaller than m. Therefore, we are able to reduce the number of linear constraints significantly to obtain a simpler problem. We will show that the two problems are closely related under several choices of T such as sub-Gaussian random projection and randomized orthorgonal system (ROS).

In this chapter, we also discuss the noised version of this problem. Instead of requiring $Ax = b$, we require that they are close enough to each other. In particular, we replace it by the condition $\|Ax - b\| \leq \delta$ for some $\delta > 0$. In this way we can avoid the assumption $m \leq n$, and we are able to assume that m is very large regardless how small the dimension n is.

4.1.1 Our contributions

The results in this chapter is inspired by the work of Mert Pilanci and Martin J. Wainwright ([22]) on the Constrained Linear Regression problem. In that work, they propose to approximate

$$x^* \in \arg\min_{x \in \mathcal{C}} \|Ax - b\|$$

by a "sketched" solution

$$\hat{x} = \arg\min_{x \in \mathcal{C}} \|TAx - Tb\|.$$

In particular, they provide bounds for the dimension d of the projected space to ensure that $\|A\hat{x} - b\| \leq (1 + \epsilon)\|Ax^* - b\|$ for any given $\epsilon > 0$. Therefore, if the original CRLM is feasible, then $\|Ax^* - b\| = 0$ and it implies that $A\hat{x} = b$. In other words, $\hat{x}$ provides a feasible solution for the original problem. In order to apply these results to our feasibilty setting, we first solve the projected problem, obtain a feasible solution $\hat{x}$ and plug it back to check if $A\hat{x} = b$ or not.

In this chapter, we are interested in the related but different question: what are the relations between the optimal objective functions of these two problems. We will prove that we can choose appropriate random projection T such that one objective value is the approximation of the other. This means that we no longer need to plug $\hat{x}$ back to validate the system $A\hat{x} = b$ (or more generally $\|A\hat{x} - b\| \leq \delta$) but can make inference immediately from the projected problem.

This result can be interesting in the case when the access to the original data is limited or unavailable. For example, for private security, the user information is strictly confidential.

Random projections provide a way to encode the problem without leaking the data of any particular people. The ability to make a decision based only on TA, Tb therefore become crucial.

Now, if the original problem is feasible, i.e. $\{x \in \mathcal{C} \mid Ax = b\}$ is not empty, then there is $\tilde{x} \in \mathcal{C}$ such that $A\tilde{x} = b$. It follows that $TA\tilde{x} = Tb$ and the projected problem is also feasible. Therefore, we only consider the infeasible case: $\{x \in \mathcal{C} \mid Ax = b\}$ is empty. In particular, we assume that

$$\|Ax^* - b\| = \min_{x \in \mathcal{C}} \|Ax - b\| > 0.$$

We will find random projection T so that the projected problem is also infeasible, i.e.

$$\min_{x \in \mathcal{C}} \|TAx - Tb\| > 0.$$

We consider two classes of random projections: σ-subgaussian matrices (see Section 2.3.1) and randomized orthogonal systems (ROS, see Section 2.3.2). For σ-subgaussian matrices, we establish the relation between the two problems via the well-known concept of Gaussian width. For a given set $\mathcal{Y} \subseteq \mathbb{R}^n$, the *Gaussian width* of $\mathcal{Y}$, denoted by $\mathbb{W}(\mathcal{Y})$ is defined by

$$\mathbb{W}(\mathcal{Y}) = \mathbb{E}_g\left(\sup_{y \in \mathcal{Y}} |\langle g, y\rangle|\right)$$

where g is of $\mathcal{N}(0, \mathbf{I})$ distribution. Gaussian width is a very important tool used in compressed sensing [7]. Our main result in this case is:

Theorem 4.1.1. *Assume that the set $\{x \in \mathcal{C} |\ Ax = b\}$ is empty. Denote by $\mathcal{Y} = A\mathcal{K} \cap \mathbb{S}^{n-1}$, in which $\mathcal{K}$ is the tangent cone of the constraint set $\mathcal{C}$ at an optimum*

$$x^* = \arg\min_{x \in \mathcal{C}} \|Ax - b\|.$$

Let $T : \mathbb{R}^m \to \mathbb{R}^k$ be a σ-subgaussian random projection. Then for any $\delta > 0$, with probability at least $1 - 6e^{-\frac{c_2 k\delta^2}{\sigma^4}}$, the set $\{x \in \mathcal{C} |\ TAx = Tb\}$ is also empty if

$$k > \left(\frac{25c_1}{1 - 17\delta}\right)^2 \mathbb{W}^2(\mathcal{Y}).$$

In the ROS case, our main result involve two generalized concept: Rademacher width and $T-$Gaussian width, which are defined respectively as follows. For any set $\mathcal{Y} \subset \mathbb{S}^{m-1}$:

$$\mathbb{R}(\mathcal{Y}) = \mathbb{E}_\varepsilon\left(\sup_{y \in \mathcal{Y}} |\langle \varepsilon, y\rangle|\right),$$

where $\varepsilon \in \{-1, +1\}^n$ is an i.i.d. vector of Rademacher variables; and

$$\mathbb{W}_T(\mathcal{Y}) = \mathbb{E}_{g,T}\left(\sup_{z \in \mathcal{Y}} |\langle g, \frac{Tz}{\sqrt{m}}\rangle|\right).$$

Using these two concepts, we obtain the following result:

Theorem 4.1.2. *Assume that the set $\{x \in \mathcal{C} |\ Ax = b\}$ is empty. Denote by $\mathcal{Y} = A\mathcal{K} \cap \mathbb{S}^{n-1}$, in which $\mathcal{K}$ is the tangent cone of the constraint set $\mathcal{C}$ at an optimum*

$$x^* = \arg\min_{x \in \mathcal{C}} \|Ax - b\|.$$

Let $T : \mathbb{R}^m \to \mathbb{R}^k$ be an ROS random projection. Then for any $\delta > 0$, with probability at least $1 - 6\left(\frac{c_1}{(mn)^2} + e^{-\frac{c_2 m\delta^2}{\mathbb{R}^2(\mathcal{Y}) + \log(mn)}}\right)$, the set $\{x \in \mathcal{C} |\ TAx = Tb\}$ is also empty if

$$\sqrt{k} > \left(\frac{736}{1 - \frac{17}{2}\delta}\right)\left(\mathbb{R}(\mathcal{Y}) + \sqrt{6\log(mn)}\right)\mathbb{W}_T(\mathcal{Y}).$$

The rest of this chapter is constructed as follows: Section 4.2 and 4.3 will be devoted to the proofs of Theorem 4.1.1 and Theorem 4.1.2. In Section 4.4, we discuss a new method to find a certificate for this system.

4.2 Random σ-subgaussian sketches

Recall from Chapter 2 that, a real-valued random variable X is said to be σ-subgaussian if one has

$$\mathbb{E}(e^{tX}) \leq e^{\frac{\sigma^2 t^2}{2}} \qquad \text{for every } t \in \mathbb{R}.$$

Now, let $T = (t_{ij}) \in \mathbb{R}^{m \times n}$ be a matrix with i.i.d entries sampling from a zero-mean σ-subgaussian distribution and $\mathrm{Var}(s_{ij}) = \frac{1}{m}$. Such a matrix will be called a *random σ-subgaussian sketch.*

Denote by $Q = T^\top T - I_{n \times n}$. From [22] we obtain the following important results:

Lemma 4.2.1. *There are universal constants c_1, c_2 such that for any $\mathcal{Y} \subseteq \mathbb{S}^{n-1}$ and any $v \in \mathbb{S}^{n-1}$, we have*

$$\sup_{u \in \mathcal{Y}} \left|u^T Q u\right| \leq c_1 \frac{\mathbb{W}(\mathcal{Y})}{\sqrt{m}} + \delta \quad \textit{with probability at least } 1 - e^{-\frac{c_2 m\delta^2}{\sigma^4}}. \tag{4.1}$$

$$\sup_{u \in \mathcal{Y}} \left|u^T Q v\right| \leq 5c_1 \frac{\mathbb{W}(\mathcal{Y})}{\sqrt{m}} + 3\delta \quad \textit{with probability at least } 1 - 3e^{-\frac{c_2 m\delta^2}{\sigma^4}}. \tag{4.2}$$

Now we will apply this lemma to prove Theorem 4.1.1.

Proof of Theorem 1:

Denote by $e := \hat{x} - x^*$. Then e belongs to the tangent cone $\mathcal{K}$ of the constraint set $\mathcal{C}$ at the optimum x^*.

Since x^* is the minimizer of $\min\{\|Ax-b\| : x \in \mathcal{C}\}$, then

$$\|Ax^*-b\|^2 \le \|A\hat{x}-b\|^2 = \|Ax^*-b\|^2 + 2\langle Ax^*-b, Ae\rangle + \|Ae\|^2.$$

Therefore, $2\langle Ax^*-b, Ae\rangle + \|Ae\|^2 \ge 0$. On the other hand, since $\hat{x}$ is the minimizer of $\min\{\|TAx-Tb\| : x \in \mathcal{C}\}$, then

$$\|TAx^*-Tb\|^2 \ge \|TA\hat{x}-Tb\|^2 = \|TAx^*-Tb\|^2 + 2\langle TAx^*-Tb, TAe\rangle + \|TAe\|^2.$$

Therefore, $\|TAe\|^2 \le -2\langle TAx^*-Tb, TAe\rangle \le 2\|TAx^*-Tb\|.\|TAe\|$, which implies that

$$\|TAe\| \le 2\|TAx^*-Tb\|.$$

Now we have

$$\begin{aligned}
\|TA\hat{x}-Tb\|^2 &= \|TAx^*-Tb\|^2 + 2\langle TAx^*-Tb, TAe\rangle + \|TAe\|^2 \\
&= \|TAx^*-Tb\|^2 + 2\langle Ax^*-b, Ae\rangle + \|Ae\|^2 + 2\langle Ax^*-b, QAe\rangle + \langle Ae, QAe\rangle \\
&\ge \|TAx^*-Tb\|^2 + 2\langle Ax^*-b, QAe\rangle + \langle Ae, QAe\rangle \\
&\ge \|TAx^*-Tb\|^2 - 2\|Ax^*-b\|.\|Ae\|\big(5c_1\frac{\mathbb{W}(\mathcal{Y})}{\sqrt{m}} + 3\delta\big) - \|Ae\|^2\big(c_1\frac{\mathbb{W}(\mathcal{Y})}{\sqrt{m}} + \delta\big)
\end{aligned}$$

with probability at least $1-4e^{-\frac{c_2 m\delta^2}{\sigma^4}}$ (by applying Lemma 4.2.1). For some arbitrary $\lambda > 0$ (we will fix it later), we have $2\|Ax^*-b\|.\|Ae\| \le \lambda\|Ax^*-b\|^2 + \frac{1}{\lambda}\|Ae\|^2$. Denote by $\Delta_1 = 5c_1\frac{\mathbb{W}(\mathcal{Y})}{\sqrt{m}} + 3\delta$ and $\Delta_2 = c_1\frac{\mathbb{W}(\mathcal{Y})}{\sqrt{m}} + \delta$, and substitute them to the above expression, we have

$$\|TA\hat{x}-Tb\|^2 \ge \|TAx^*-Tb\|^2 - \lambda\Delta_1\|Ax^*-b\|^2 - \big(\frac{1}{\lambda}\Delta_1 + \Delta_2\big)\|Ae\|^2 \tag{4.3}$$

with probability at least $1-4e^{-\frac{c_2 m\delta^2}{\sigma^4}}$. Now, let $y^* = \frac{Ax^*-b}{\|Ax^*-b\|}$, we have

$$\begin{aligned}
\|TAx^*-Tb\|^2 &= \|Ax^*-b\|^2 + \langle Ax^*-b, Q(Ax^*-b)\rangle \\
&\ge \|Ax^*-b\|^2\big(1 - c_1\frac{\mathbb{W}(\{y^*\})}{\sqrt{m}} - \delta\big) \\
&\ge (1-\Delta_2)\|Ax^*-b\|^2
\end{aligned} \tag{4.4}$$

with probability at least $1-e^{-\frac{c_2 m\delta^2}{\sigma^4}}$. Here, the first inequality is an direct application of Lemma 4.2.1 and the second inequality follows from the fact that $\mathbb{W}(\{y^*\}) \le \mathbb{W}(\mathcal{Y})$.

Furthermore,

$$\|TAe\|^2 = \|Ae\|^2 + \langle Ae, Q(Ae)\rangle \ge \|Ae\|^2\big(1 - c_1\frac{\mathbb{W}(\mathcal{Y})}{\sqrt{m}} - \delta\big) = (1-\Delta_2)\|Ae\|^2,$$

which implies

$$\|Ae\|^2 \leq \frac{1}{1-\Delta_2}\|TAe\|^2 \leq \frac{4}{1-\Delta_2}\|TAx^* - Tb\|^2 \tag{4.5}$$

with probability at least $1 - e^{-\frac{c_2 m\delta^2}{\sigma^4}}$ (due to the fact that $\|TAe\| \leq 2\|TAx^* - Tb\|$).

From (4.3), (4.4) and (4.5) we have

$$\begin{aligned}
\|TA\hat{x} - Tb\|^2 &\geq \|TAx^* - Tb\|^2 - \lambda\Delta_1\|Ax^* - b\|^2 - 4\frac{\frac{1}{\lambda}\Delta_1 + \Delta_2}{1-\Delta_2}\|TAx^* - Tb\|^2 \\
&= \left(1 - 4\frac{\frac{1}{\lambda}\Delta_1 + \Delta_2}{1-\Delta_2}\right)\|TAx^* - Tb\|^2 - \lambda\Delta_1\|Ax^* - b\|^2 \\
&\geq \big(1 - 4\frac{\frac{1}{\lambda}\Delta_1 + \Delta_2}{1-\Delta_2}\big)(1-\Delta_2)\|Ax^* - b\|^2 - \lambda\Delta_1\|Ax^* - b\|^2 \\
&= (1 - \Delta_2 - \frac{4}{\lambda}\Delta_1 - 4\Delta_2 - \lambda\Delta_1)\|Ax^* - b\|^2
\end{aligned}$$

with probability at least $1 - 6e^{-\frac{c_2 m\delta^2}{\sigma^4}}$. Select (the best possible) $\lambda = 2$, then

$$\|TA\hat{x} - Tb\|^2 \geq (1 - 4\Delta_1 - 5\Delta_2)\|Ax^* - b\|^2 = (1 - 25c_1\frac{\mathbb{W}(\mathcal{Y})}{\sqrt{m}} - 17\delta)\|Ax^* - b\|^2$$

with probability at least $1 - 6e^{-\frac{c_2 m\delta^2}{\sigma^4}}$.

It follows that $TA\hat{x} \neq Tb$ with probability at least $1 - 6e^{-\frac{c_2 m\delta^2}{\sigma^4}}$, if

$$m > \big(\frac{25c_1}{1-17\delta}\big)^2\mathbb{W}^2(\mathcal{Y}),$$

which proves Theorem 4.1.1. □

4.3 Random orthonormal systems

Sketched matrices resulted from σ-subgaussian distributions are often dense, therefore requiring us to perform expensive matrix-vector multiplications. In order to overcome that difficulty, a different kind of sketch matrices is proposed in [4]. The idea is to randomly sample rows from an orthogonal matrix to form the new matrix T. Formally, we define this matrix as follows: let H be an orthogonal matrix, we independently sample each row of T by

$$s_i = \sqrt{n}DH^T p_i,$$

where p_i is a vector chosen uniformly at random from the set of all n canonical basis vectors $\{e_1, \ldots, e_n\}$, and $D = \text{diag}\,(v)$ is a diagonal matrix of i.i.d Rademacher variables $v \in \{-1, +1\}^n$. It is showed that if for certain choices of H (such as Walsh - Hadamard

matrix), the matrix-vector product Tx can be computed in $O(n \log m)$ time for any $x \in \mathbb{R}^n$. This is a huge save of time, comparing to the normal $O(nm)$ time from an unstructured matrix.

For simplicity, we define a function Φ given by:

$$\Phi(\mathcal{X}) = 8\big\{\mathbb{R}(\mathcal{X}) + \sqrt{6\log(mn)}\big\}\frac{\mathbb{W}_T(\mathcal{X})}{\sqrt{m}}.$$

Then from [22], we have the following important properties (note that the parameters in this lemma have been corrected based on the original proof):

Lemma 4.3.1. *There are universal constants c_1, c_2 such that for any $\mathcal{Y} \subseteq \mathbb{B}^n$ and any $v \in \mathbb{S}^{n-1}$, we have*

$$\sup_{u\in\mathcal{Y}} \big|u^T Q u\big| \le \Phi(\mathcal{Y}) + \frac{\delta}{2} \qquad \text{with probability at least } 1 - c_1\tfrac{1}{(mn)^2} - c_1 e^{-\frac{c_2 m\delta^2}{\mathbb{R}^2(\mathcal{Y})+\log(mn)}}. \quad (4.6)$$

$$\sup_{u\in\mathcal{Y}} \big|u^T Q v\big| \le 21\Phi(\mathcal{Y}) + \frac{3\delta}{2} \qquad \text{with probability at least } 1 - 3c_1\tfrac{1}{(mn)^2} - 3c_1 e^{-\frac{c_2 m\delta^2}{\mathbb{R}^2(\mathcal{Y})+\log(mn)}}. \quad (4.7)$$

Proof of Theorem 4.1.2:

This proof follows almost the same line as that of Theorem 4.1.1. Denote by $e := \hat{x} - x^*$. Then similar to the previous proof, we also have:

$$2\langle Ax^* - b, Ae\rangle + \|Ae\|^2 \ge 0 \quad (4.8)$$

$$\|TAe\| \le 2\|TAx^* - Tb\|. \quad (4.9)$$

Then we have

$$\begin{aligned}
\|TA\hat{x} - Tb\|^2 &= \|TAx^* - Tb\|^2 + 2\langle TAx^* - Tb, TAe\rangle + \|TAe\|^2 \\
&= \|TAx^* - Tb\|^2 + 2\langle Ax^* - b, Ae\rangle + \|Ae\|^2 + 2\langle Ax^* - b, QAe\rangle + \langle Ae, QAe\rangle \\
&\ge \|TAx^* - Tb\|^2 + 2\langle Ax^* - b, QAe\rangle + \langle Ae, QAe\rangle \\
&\ge \|TAx^* - Tb\|^2 - 2\|Ax^* - b\|.\|Ae\|\big(21\,\Phi(\mathcal{Y}) + \frac{3\delta}{2}\big) - \|Ae\|^2\big(\Phi(\mathcal{Y}) + \frac{\delta}{2}\big)
\end{aligned}$$

with probability at least $1 - 4\left(\frac{c_1}{(mn)^2} + c_1 e^{-\frac{c_2 m\delta^2}{\mathbb{R}^2(\mathcal{Y})+\log(mn)}}\right)$ (by applying Lemma 4.3.1). For some arbitrary $\lambda > 0$ (we will fix it later), we have $2\|Ax^* - b\|.\|Ae\| \le \lambda\|Ax^* - b\|^2 + \frac{1}{\lambda}\|Ae\|^2$. Denote by $\Delta_1 = 21\,\Phi(\mathcal{Y}) + \frac{3\delta}{2}$ and $\Delta_2 = \Phi(\mathcal{Y}) + \frac{\delta}{2}$, and substitute them to the above expression, then we have

$$\|TA\hat{x} - Tb\|^2 \ge \|TAx^* - Tb\|^2 - \lambda\Delta_1\|Ax^* - b\|^2 - \big(\frac{1}{\lambda}\Delta_1 + \Delta_2\big)\|Ae\|^2 \quad (4.10)$$

with probability at least $1-4\left(\frac{c_1}{(mn)^2}+c_1e^{-\frac{c_2m\delta^2}{\mathbb{R}^2(\mathcal{Y})+\log(mn)}}\right)$. Now, let $y^*:=\frac{Ax^*-b}{\|Ax^*-b\|}$, we have

$$\begin{aligned}\|TAx^*-Tb\|^2&=\|Ax^*-b\|^2+\langle Ax^*-b,Q(Ax^*-b)\rangle\\&\geq\|Ax^*-b\|^2\big(1-\Phi(\{y^*\})-\frac{\delta}{2}\big)\\&\geq(1-4\Phi(\mathcal{Y})-\frac{\delta}{2})\|Ax^*-b\|^2\end{aligned}\tag{4.11}$$

with probability at least $1-\left(\frac{c_1}{(mn)^2}+c_1e^{-\frac{c_2m\delta^2}{\mathbb{R}^2(\mathcal{Y})+\log(mn)}}\right)$. Here we use the fact (from the proof of Lemma 5 in [22]) that $\Phi(\{y^*\})\leq 4\Phi(\mathcal{Y})$.

Similar to the previous proof, we also have

$$\|TAe\|^2=\|Ae\|^2+\langle Ae,Q(Ae)\rangle\geq(1-\Delta_2)\|Ae\|^2,$$

therefore

$$\|Ae\|^2\leq\frac{1}{1-\Delta_2}\|TAe\|^2\leq\frac{4}{1-\Delta_2}\|TAx^*-Tb\|^2\tag{4.12}$$

with probability at least $1-\left(\frac{c_1}{(mn)^2}+c_1e^{-\frac{c_2m\delta^2}{\mathbb{R}^2(\mathcal{Y})+\log(mn)}}\right)$.

From (4.10), (4.11) and (4.12) we have

$$\begin{aligned}\|TA\hat{x}-Tb\|^2&\geq\|TAx^*-Tb\|^2-\lambda\Delta_1\|Ax^*-b\|^2-4\frac{\frac{1}{\lambda}\Delta_1+\Delta_2}{1-\Delta_2}\|TAx^*-Tb\|^2\\&=\left(1-4\frac{\frac{1}{\lambda}\Delta_1+\Delta_2}{1-\Delta_2}\right)\|TAx^*-Tb\|^2-\lambda\Delta_1\|Ax^*-b\|^2\\&\geq\big(1-4\frac{\frac{1}{\lambda}\Delta_1+\Delta_2}{1-\Delta_2}\big)\big(1-4\Phi(\mathcal{Y})-\frac{\delta}{2}\big)\|Ax^*-b\|^2-\lambda\Delta_1\|Ax^*-b\|^2\\&\geq\left(1-4\Phi(\mathcal{Y})-\frac{\delta}{2}-\frac{4}{\lambda}\Delta_1-4\Delta_2-\lambda\Delta_1\right)\|Ax^*-b\|^2\end{aligned}$$

with probability at least $1-4\left(\frac{c_1}{(mn)^2}+e^{-\frac{c_2m\delta^2}{\mathbb{R}^2(\mathcal{Y})+\log(mn)}}\right)$. (Here we simplify by using the fact that $1-4\Phi(\mathcal{Y})-\frac{\delta}{2}<1-\Delta_2$). Now select $\lambda=2$ (the best possible), then we have

$$\begin{aligned}\|TA\hat{x}-Tb\|^2&\geq(1-4\Phi(\mathcal{Y})-\frac{\delta}{2}-4\Delta_1-4\Delta_2)\|Ax^*-b\|^2\\&=(1-92\Phi(\mathcal{Y})-\frac{17}{2}\delta)\|Ax^*-b\|^2\end{aligned}$$

with probability at least $1-6\left(\frac{c_1}{(mn)^2}+e^{-\frac{c_2m\delta^2}{\mathbb{R}^2(\mathcal{Y})+\log(mn)}}\right)$.

The last expression is positive if and only if

$$1 - \frac{17}{2}\delta - 736\big(\mathbb{R}(\mathcal{Y}) + \sqrt{6\log(mn)}\big)\frac{\mathbb{W}_T(\mathcal{Y})}{\sqrt{m}} > 0,$$

or equivalently,

$$\sqrt{m} > \left(\frac{736}{1 - \frac{17}{2}\delta}\right)\big(\mathbb{R}(\mathcal{Y}) + \sqrt{6\log(mn)}\big)\mathbb{W}_T(\mathcal{Y}),$$

which proves Theorem 4.1.2.

□

4.4 Sketch-and-project method

Assume that the CRLM problem we considered in this chapter is feasible. As shown in Chapter 3, it is not possible to find a certificate for this system just by solving the projected problem. In this section, we will propose a method for finding such a certificate. The idea is to first sketch the constrained system by a random projection, and then to project a current solution to the new sketched feasible set.

Algorithm 1 Randomized sketch-and-project (RSP)

1: **parameter:** $\mathcal{D}$ = distribution over random matrices
2: Choose $x^0 \in \mathbb{R}^n$ ▷ Initialization step
3: **for** $k = 0, 1, 2, \ldots$ **do**
4: Construct a random projection matrix $S \sim \mathcal{D}$
5: $x^{k+1} = \arg\min \|x - x^k\|^2$ s.t. $SAx = Sb, x \in C$. ▷ Update step

For any vector x and any set Q, we will denote by

$$\|x - Q\| = \min_{y \in Q} \|x - y\|.$$

Let $P = \{x \in C \mid Ax = b\}$ and $P_S = \{x \in C \mid SAx = Sb\}$.

Then we can rewrite the update step simply as

$$x^{k+1} = \arg\min_{x \in P_S} \|x - x^k\|^2.$$

In other word, $\|x^k - x^{k+1}\| = \|x^k - P_S\|$.

We have the following result:

Lemma 4.4.1. *Given x^k, x^{k+1}, P, P_S defined as above, then*

(i) For any $x \in P$, $\langle x^k - x^{k+1}, x - x^{k+1}\rangle \leq 0$.

(ii) $\|x^{k+1} - P\|^2 \leq \|x^k - P\|^2 - \|x^k - P_S\|^2$.

Proof. (i) Assume there is some $\hat{x} \in P$ such that $\langle x^k - x^{k+1}, \hat{x} - x^{k+1}\rangle > 0$. For any $\lambda \in (0,1)$, let define $x_\lambda = \lambda\hat{x} + (1-\lambda)x^{k+1}$. Since $P \subseteq P_S$, $\hat{x} \in P_S$. Then by the convexity of P_S (followed from the convexity of C), we have $x_\lambda \in P_S$. However,

$$\begin{aligned}\|x^k - x_\lambda\|^2 &= \|x^k - x^{k+1} + \lambda(x^{k+1} - \hat{x})\|^2 \\ &= \|x^k - x^{k+1}\|^2 + \lambda^2\|x^{k+1} - \hat{x}\|^2 - 2\lambda\langle x^k - x^{k+1}, \hat{x} - x^{k+1}\rangle \\ &< \|x^k - x^{k+1}\|^2 \text{ when } \lambda \text{ is small enough.}\end{aligned}$$

But this contradicts to the definition of x^{k+1}.
(ii) Let denote by

$$x_k^* = \arg\min_{x \in P} \|x - x^k\|^2,$$

then

$$\begin{aligned}\|x^k - P\|^2 &= \|x^k - x_k^*\|^2 \\ &= \|x^k - x^{k+1} + x^{k+1} - x_k^*\|^2 \\ &= \|x^k - x^{k+1}\|^2 + \|x^{k+1} - x_k^*\|^2 + 2\langle x^k - x^{k+1}, x^{k+1} - x_k^*\rangle \\ &\geq \|x^k - x^{k+1}\|^2 + \|x^{k+1} - x_k^*\|^2 \qquad \text{(due to (i))} \\ &= \|x^k - P_S\|^2 + \|x^{k+1} - x_k^*\|^2 \\ &\geq \|x^k - P_S\|^2 + \|x^{k+1} - P\|^2,\end{aligned}$$

which proves (ii). □ □

Now, we are going to estimate the quantity $\|x^k - P_S\|^2$ for some special construction of the random matrix S. We will consider the case when the entries of S are i.i.d random variables sampled from a σ-subgaussian distribution.

Proposition 4.4.2. *There are universal constants c_1, c_2 such that, with probability at least $1 - e^{-\frac{c_2 d\delta^2}{\sigma^4}}$ we have*

$$\|x^{k+1} - P\|^2 \leq \big(\frac{\mathbb{W}(A\mathcal{K})}{\sqrt{m}} + \delta\big)\|x^k - P\|^2,$$

Denote by $Q := S^TS - E$ where E is the identity matrix. Then $S^TS = Q + E$.

Proof. By the definition, we have

$$\|x^k - P_S\|^2 = \min_{x \in C,\, SAx = Sb} \|x^k - x\|^2.$$

Notice that, by using a penalty term, it can also be written as

$$\max_{\lambda\geq 0}\left(\min_{x\in C}\|x^k-x\|^2+\lambda\|SAx-Sb\|^2\right). \tag{4.13}$$

Indeed, if we denote by $L_\lambda=\min_{x\in C}\|x^k-x\|^2+\lambda\|SAx-Sb\|^2$ for each $\lambda\geq 0$. Then due to $x^{k+1}\in C$, we have

$$L_\lambda\leq\|x^k-x^{k+1}\|^2+\lambda\|SAx^{k+1}-Sb\|^2=\|x^k-P_S\|^2,$$

which proves that the value in (4.13) is smaller than or equal to $\|x^k-P_S\|^2$. But when $\lambda\geq 0$ is very large, the penalty becomes so expensive that it forces $SAx=Sb$, and in those cases, $L_\lambda=\|x^k-P_S\|^2$. The claim is proved.

Now, for any $\lambda\geq 0$, we have,

$$\begin{aligned}
L_\lambda&=\min_{x\in C}\|x^k-x\|^2+\lambda\|SAx-Sb\|^2\\
&=\min_{x\in C}\|x^k-x\|^2+\lambda\langle S(Ax-b),S(Ax-b)\rangle\\
&=\min_{x\in C}\|x^k-x\|^2+\lambda\langle S^TS(Ax-b),Ax-b\rangle\\
&=\min_{x\in C}\|x^k-x\|^2++\lambda\|Ax-b\|^2+\lambda\langle Q(Ax-b),Ax-b\rangle.\\
&=\min_{x\in C}\|x^k-x\|^2++\lambda\|Ax-b\|^2-\lambda\|Ax-b\|^2\big(\frac{\mathbb{W}(A\mathcal{K})}{\sqrt{m}}+\delta\big)
\end{aligned}$$

with probability at least $1-\delta$. Therefore, we have

$$L_\lambda\geq\big(1-\frac{\mathbb{W}(A\mathcal{K})}{\sqrt{m}}-\delta\big)\min_{x\in C}\|x^k-x\|^2+\lambda\|Ax-b\|^2$$

Therefore, we have

$$\begin{aligned}
\|x^k-P_S\|^2&=\max_{\lambda\geq 0}\ L_\lambda\\
&\geq\big(1-\frac{\mathbb{W}(A\mathcal{K})}{\sqrt{m}}-\delta\big)\max_{\lambda\geq 0}\left(\min_{x\in C}\|x^k-x\|^2+\lambda\|Ax-b\|^2\right)\\
&=\big(1-\frac{\mathbb{W}(A\mathcal{K})}{\sqrt{m}}-\delta\big)\min_{x\in C,\ Ax=b}\|x^k-x\|^2\\
&=\big(1-\frac{\mathbb{W}(A\mathcal{K})}{\sqrt{m}}-\delta\big)\|x^k-P\|^2
\end{aligned}$$

with probability at least $1-\delta$. Therefore, we have, from part (ii) of Lemma ,

$$\|x^{k+1}-P\|^2\leq\|x^k-P\|^2-\|x^k-P_S\|^2\leq\big(\frac{\mathbb{W}(A\mathcal{K})}{\sqrt{m}}+\delta\big)\|x^k-P\|^2,$$

with probability at least $1-\delta$.

□

Chapter 5

Gaussian random projections for general membership problems

5.1 Introduction

In this chapter we employ random projections to study the following general problem:

> Euclidean Set Membership Problem (ESMP). Given $b \in \mathbb{R}^m$ and $S \subseteq \mathbb{R}^m$, decide whether $b \in S$.

This is a generalization of the convex hull and cone membership problems that we consider in Chapter 2. However, in this chapter, we consider the general case where the data set S has no specific structure. We will use Gaussian random projections in our arguments to embed both b and S to a lower dimensional space, and solve its associated projected version:

> Projected ESMP (PESMP). Given $b \in \mathbb{R}^m$ and $S \subseteq \mathbb{R}^m$, and let $T : \mathbb{R}^m \to \mathbb{R}^d$ be some given mapping. Decide whether $T(b) \in T(S)$.

Our objective is to investigate the relationships between ESMP and PESMP. As before, it is obvious that when $b \in S$ then $T(b) \in T(S)$. We are therefore only interested in the case when $b \notin S$, i.e. we want to estimate $\mathsf{Prob}(T(b) \notin T(S))$, given that $b \notin S$.

5.1.1 Our contributions

In the case when S is at most countable (i.e. finite or countable), using a straightforward argument, we prove that these two problems are equivalent almost surely. However, this result

is only of theoretical interest due to round-off errors in floating point operations, which make its practical application difficult. We address this issue by introducing a threshold $\tau > 0$ and study the corresponding "thresholded" problem:

> THRESHOLD ESMP (TESMP): Given b, S, T as above. Let $\delta > 0$. Decide whether $\|T(b) - T(S)\| \geq \tau$.

In the case when S may also be uncountable, we prove that ESMP and PESMP are also equivalent if the projected dimension d is proportional to some intrinsic dimension of the set S. In particular, we employ the definition of doubling dimension (defined later) to prove that, if $b \notin S$, then $T(b) \notin T(S)$ almost surely as long as the projected dimension $d \geq c \cdot ddim(S)$. Here, $ddim(S)$ is the doubling dimension of S and c is some universal constant.

We also extend this result to the threshold case, and obtain a more useful bound for d.

5.2 Finite and countable sets

In this section, we assume that S is either finite or countable. Let $T \in \mathbb{R}^{d\times m}$ be a random matrix drawn from Gaussian distribution, i.e. each entry of T is independently sampled from $\mathcal{N}(0,1)$. It is well known that, for an arbitrary unit vector $a \in \mathbb{S}^{m-1}$, then $\|Ta\|^2$ is a random variable with a Chi-squared distribution χ^2_d with d degrees of freedom ([20]). Its corresponding density function is $\frac{2^{-d/2}}{\Gamma(d/2)}x^{d/2-1}e^{d/2}$, where $\Gamma(\cdot)$ is the gamma function. By [9], for any $0 < \delta < 1$, taking $z = \frac{\delta}{d}$ yields a cumulative distribution function (CDF)

$$F_{\chi^2_d}(\delta) \leq (ze^{1-z})^{d/2} < (ze)^{d/2} = \left(\frac{e\delta}{d}\right)^{d/2}. \tag{5.1}$$

Thus, we have

$$\mathsf{Prob}(\|Ta\| \leq \delta) = F_{\chi^2_d}(\delta^2) < (\frac{3}{d}\delta^2)^{d/2} \tag{5.2}$$

or, more simply, $\mathsf{Prob}(\|Ta\| \leq \delta) < \delta^d$ when $d \geq 3$.

Using this estimation, we immediately obtain the following result.

Proposition 5.2.1. *Given $b \in \mathbb{R}^m$ and $S \subseteq \mathbb{R}^m$, at most countable, such that $b \notin S$. Then, for any $d \geq 1$ and for a Gaussian random projection $T : \mathbb{R}^m \to \mathbb{R}^d$, we have $T(b) \notin T(S)$ almost surely, i.e. $\mathsf{Prob}\big(T(b) \notin T(S)\big) = 1$.*

Proof. First, note that for any $u \neq 0$, $Tu \neq 0$ holds almost certainly. Indeed, without loss of generality we can assume that $\|u\| = 1$. Then for any $0 < \delta < 1$:

$$\mathsf{Prob}\big(T(u) = 0\big) \leq \mathsf{Prob}\big(\|Tu\| \leq \delta\big) < (3\delta^2)^{d/2} \to 0 \text{ as } \delta \to 0. \tag{5.3}$$

It means that for any $y \neq b$, then $T(y) \neq T(b)$ almost surely. Since the event $T(b) \notin T(S)$ can be written as the intersection of at most countably many "almost sure" events $T(b) \neq T(y)$ (for $y \in S$), it follows that $\mathsf{Prob}\big(T(b) \notin T(S)\big) = 1$, as claimed. □

Proposition 5.2.1 is simple, but it looks interesting because it suggests that we only need to project the data points to a line (i.e. $d = 1$) and study an equivalent membership problem on a line. Furthermore, it turns out that this result remains true for a large class of random projections.

Proposition 5.2.2. *Let ν be a probability distribution on $\mathbb{R}^m$ with bounded Lebesgue density f. Let $S \subseteq \mathbb{R}^m$ be an at most countable set such that $0 \notin S$. Then, for a random projection $T : \mathbb{R}^m \to \mathbb{R}^1$ sampled from ν, we have $0 \notin T(S)$ almost surely, i.e. $\mathsf{Prob}\big(0 \notin T(S)\big) = 1$.*

Proof. For any $0 \neq s \in Y$, consider the set $\mathcal{E}_s = \{T : \mathbb{R}^m \to \mathbb{R}^1 \mid T(s) = 0\}$. If we regard each $T : \mathbb{R}^m \to \mathbb{R}^1$ as a vector $t \in \mathbb{R}^m$, then $\mathcal{E}_s$ is a hyperplane $\{t \in \mathbb{R}^m | \ s \cdot t = 0\}$ and we have

$$\mathsf{Prob}(T(s) = 0) = \nu(\mathcal{E}_s) = \int_{\mathcal{E}_s} f d\mu \leq \|f\|_\infty \int_{\mathcal{E}_s} d\mu = 0$$

where μ denotes the Lebesgue measure on $\mathbb{R}^m$. The proof then follows by the countability of S, similarly to Proposition 5.2.1. □

This idea, however, does not work in practice: we tested it by considering the ESMP given by the IPF defined on the set $\{x \in \mathbb{Z}_+^n \cap [L, U] \mid Ax = b\}$. Numerical experiments indicate that the corresponding PESMP $\{x \in \mathbb{Z}_+^n \cap [L, U] \mid T(A)x = T(b)\}$, with T consisting of a one-row Gaussian projection matrix, is always feasible despite the infeasibility of the original IPF. Since Prop. 5.2.1 assumes that the components of T are real numbers, the reason behind this failure is possibly due to the round-off error associated to the floating point representation used in computers. Specifically, when $T(A)x$ is too close to $T(b)$, floating point operations will consider them as a single point. In order to address this issue, we force the projected problems to obey stricter requirements. In particular, instead of only requiring that $T(b) \notin T(S)$, we ensure that

$$\mathsf{dist}(T(b), T(S)) = \min_{x \in S} \|T(b) - T(x)\| > \tau, \tag{5.4}$$

where dist denotes the Euclidean distance, and $\tau > 0$ is a (small) given constant. With this restriction, we obtain the following result.

Proposition 5.2.3. *Given $\tau > 0$, $0 < \delta < 1$ and $b \notin S \subseteq \mathbb{R}^m$, where S is a finite set. Let $R = \min_{x \in S} \|b - x\| > 0$ and $T : \mathbb{R}^m \to \mathbb{R}^d$ be a Gaussian random projection with $d \geq \frac{\log(\frac{|S|}{\delta})}{\log(\frac{R}{\tau})}$. Then:*

$$\mathsf{Prob}\big(\min_{x \in S} \|T(b) - T(x)\| > \tau\big) > 1 - \delta.$$

Proof. We assume that $d \geq 3$. For any $x \in S$, by the linearity of T, we have:

$$\begin{aligned}\mathsf{Prob}\big(\|T(b-x)\| \leq \tau\big) &= \mathsf{Prob}\bigg(\Big\|T\big(\frac{b-x}{\|b-x\|}\big)\Big\| \leq \frac{\tau}{\|b-x\|}\bigg) \\ &\leq \mathsf{Prob}\bigg(\Big\|T\big(\frac{b-x}{\|b-x\|}\big)\Big\| \leq \frac{\tau}{R}\bigg) < \frac{\tau^d}{R^d},\end{aligned}$$

due to (5.1). Therefore, by the union bound,

$$\begin{aligned}\mathsf{Prob}\big(\min_{x\in S} \|T(b)-T(x)\| > \tau\big) &= 1 - \mathsf{Prob}\big(\min_{x\in S} \|T(b)-T(x)\| \leq \tau\big) \\ &\geq 1 - \sum_{x\in S} \mathsf{Prob}\big(\|T(b)-T(x)\| \leq \tau\big) \\ &> 1 - |S|\frac{\tau^d}{R^d}.\end{aligned}$$

The RHS is greater than or equal to $1-\delta$ if and only if $\frac{R^d}{\tau^d} \geq \frac{|S|}{\delta}$, which is equivalent to $d \geq \frac{\log(\frac{|S|}{\delta})}{\log(\frac{R}{\tau})}$, as claimed. □

Note that R is often unknown and can be arbitrarily small. However, if both b, S are integral and $\tau < 1$, then $R \geq 1$ and we can select $d > \frac{\log \frac{|S|}{\delta}}{\log \frac{1}{\tau}}$ in the above proposition.

5.3 Sets with low doubling dimensions

In many real-world applications, the data set S is not finite or countable, but it lies in some intrinsically low-dimensional space. Examples of such sets are various, including human motion records, speed signals, image and text data ([8, 6]). Random projections can provide a tool to extract the full information of the set, despite of the (high) dimension of the ambient space that it is embedded in.

In this section, we will use *doubling dimension* as a measure for the intrinsic dimension of a set. Let denote by $B(x,r) = \{y \in S : \|y-x\| \leq r\}$, i.e. the closed ball centered at x with radius $r > 0$ (w.r.t S). We use the following definition:

Definition 5.3.1. *The* doubling constant *of a set $S \subseteq \mathbb{R}^m$ is the smallest number λ_S such that any closed ball in S can be covered by at most λ_S closed balls of half the radius. A set $S \subseteq \mathbb{R}^m$ is called a* doubling set *if it has a finite doubling constant. The number $\log_2(\lambda_S)$ is then called the* doubling dimension *of S and denoted by $\mathsf{ddim}\ (S)$.*

One popular example of doubling spaces is a Euclidean space $\mathbb{R}^m$. In this case, it is well-known that its doubling dimension is a constant factor of m ([23, 11]). However, there are cases where the set $S \subseteq \mathbb{R}^m$ is of much lower doubling dimension. It is also easy to see that

$\mathsf{ddim}\ (S)$ does not depend on the dimension m. Therefore, doubling dimension becomes a powerful tool for reducing dimensions in several classes of problems such as nearest neighbor [15, 12], low-distortion embeddings [2], and clustering [17].

In this section, we assume that $\mathsf{ddim}\ (S)$ is relatively small compared to m. Note that, computing the doubling dimension of S is generally **NP**-hard ([10]). For simplicity, we assume that λ_S is a power of 2, i.e. the doubling dimension of S is a positive integer number.

We shall make use of the following simple lemma.

Lemma 5.3.2. *For any $b \in S$ and $\varepsilon, r > 0$, there is a set $S_0 \subseteq S$ of size at most $\lambda_S^{\lceil \log_2(\frac{r}{\varepsilon}) \rceil}$ such that*

$$B(b, r) \subseteq \bigcup_{s \in S_0} B(s, \varepsilon). \tag{5.5}$$

Proof. By the definition of the doubling dimension, $B(b, r)$ is covered by at most λ_S closed balls of radius $\frac{r}{2}$. Each of these balls in turn is covered by λ_S balls of radius $\frac{r}{4}$, and so on: iteratively, for each $k \geq 1$, $B(b, r)$ is covered by λ_S^k balls of radius $\frac{r}{2^k}$. If we select $k = \lceil \log_2(\frac{r}{\varepsilon}) \rceil$ then $k \geq \log_2(\frac{r}{\varepsilon})$, i.e. $\frac{r}{2^k} \leq \varepsilon$. This means $B(b, r)$ is covered by $\lambda_S^{\lceil \log_2(\frac{r}{\varepsilon}) \rceil}$ balls of radius ε. $\square$

We will also use the following lemma:

Lemma 5.3.3. *Let $S \subseteq \{s \in \mathbb{R}^m | \, \|s\| \leq 1\}$ be a subset of the m-dimensional Euclidean unit ball. Let $T : \mathbb{R}^m \to \mathbb{R}^d$ be a Gaussian random projection. Then there exist universal constants $c, C > 0$ such that for $d \geq C \log \lambda_S + 1$ and $\delta \geq 6$, the following holds:*

$$\mathsf{Prob}(\exists s \in S \ s.t. \ \|Ts\| > \delta) < e^{-cd\delta^2}. \tag{5.6}$$

This lemma is proved in [12] using concentration estimations for sum of squared gaussian variables (Chi-squared distribution). In particular, we recall two important inequalities: for all $\delta > 0$ and a mapping T as above,

$$\mathsf{Prob}\ (|\|Ta\| - 1| > \delta) \leq e^{-d\delta^2/8} \quad \text{and} \tag{5.7}$$

$$\mathsf{Prob}(\|Ta\| \leq \delta) < \delta^d \quad \text{when } d \geq 3 \text{ (already used in the previous section).} \tag{5.8}$$

For the sake of completeness, we will present the original proof ([12]). Here, we use an additional requirement that $\delta \geq 6$ instead of $\delta > 1$ to make the proof more rigorous (than the original), however the main argument is unchanged.

Proof. Choose $b = 0, r = 1$ and $\varepsilon_k = \frac{1}{2^k}$ in the Lemma 5.3.2. By earlier convention that $B(x, r) = \{y \in S : \|y - x\| \leq r\}$, obviously we have $S = B(0, 1)$. Then there is a set $S_k \subseteq S$

of size at most λ_S^k such that

$$S \subseteq \bigcup_{s \in S_k} B(s, \frac{1}{2^k}). \tag{5.9}$$

Therefore, for any $x \in S$, there is a sequence $\{0 = x_0, x_1, x_2, x_3, \ldots\}$ that converges to x, with $x_k \in S_k$ and $\|x_k - x_{k+1}\| \leq \frac{1}{2^k}$ for each k. We claim that, if $\|Tx\| \geq \delta$, then there must be some k such that

$$\|T(x_k - x_{k+1})\| \geq \frac{\delta}{3}(\frac{3}{2})^{-k}$$

Indeed, if no such k exists, then

$$\delta \leq \|Tx\| \leq \sum_{k=0}^{\infty} \|T(x_k - x_{k+1})\| < \frac{\delta}{3} \sum_{k=0}^{\infty} (\frac{3}{2})^{-k} = \delta, \text{ a contradiction.}$$

Now, if we want to neglect x, we can treat x_k and x_{k+1} (found above) as two points u, v, in which $u \in S_k$ and $v \in B(u, \frac{1}{2^k}) \cap S_{k+1}$. Then we have

$$\|Tu - Tv\| > \frac{\delta}{3}(\frac{3}{2})^{-k} > \frac{\|u - v\|}{2^{-k}} \frac{\delta}{3}(\frac{3}{2})^{-k} = \frac{\delta}{3}(\frac{4}{3})^k \|u - v\|.$$

Therefore, we have

$$\begin{aligned}
& \mathsf{Prob}\big(\exists x \in S \text{ s.t } \|T(x)\| > \delta\big) \\
\leq\ & \mathsf{Prob}\bigg(\exists x \in S, \exists k \geq 0 \text{ s.t } \|T(x_k - x_{k+1})\| > \frac{\delta}{3}(\frac{3}{2})^{-k}\bigg) \\
\leq\ & \sum_{k=0}^{\infty} \mathsf{Prob}\bigg(\exists u \in S_k,\ \exists v \in B(u, \frac{1}{2^k}) \cap S_{k+1} \text{ s.t } \|Tu - Tv\| > \frac{\delta}{3}(\frac{4}{3})^k \|u - v\|\bigg) \\
\leq\ & \sum_{k=0}^{\infty} \lambda_S^{2k+1} \mathsf{Prob}\bigg(\|Tz\| > \frac{\delta}{3}(\frac{4}{3})^k\bigg) \quad \text{for some unit vector } z \text{ (by the union bound).}
\end{aligned}$$

Since $\delta \geq 6$, for any $k \geq 0$ we have $\frac{\delta}{3}(\frac{4}{3})^k \geq 1 + \frac{\delta}{6}(\frac{4}{3})^k$. Therefore, from the inequality (5.7), the above expression is less than or equal to

$$\sum_{k=0}^{\infty} \lambda_S^{2k+1} \mathsf{Prob}\big(|\|Tz\| - 1| \geq \frac{\delta}{6}(\frac{4}{3})^k\big) \ \leq \ \sum_{k=0}^{\infty} \lambda_S^{2k+1} e^{-d\frac{\delta^2}{8.6^2}(\frac{4}{3})^{2k}} \ \leq \ e^{-cd\delta^2} \tag{5.10}$$

as long as $d \geq C \log(\lambda_S)$ for some universal constants c, C. The proof is done. □

Now we state one of the main results in this chapter. It says that we can maintain the equivalence between ESMP and its projected version by using Gaussian random projections with d proportional to the doubling dimension of S.

Theorem 5.3.4. *Given $b \notin S \subseteq \mathbb{R}^m$ where S is a closed doubling set. Let $T : \mathbb{R}^m \to \mathbb{R}^d$ be a Gaussian random projection. Then*

$$\mathsf{Prob}(T(b) \notin T(S)) = 1 \tag{5.11}$$

if $d \geq \mathcal{C}\,\mathsf{ddim}\,(S)$, for some universal constant $\mathcal{C}$.

Proof. Let $d \geq \mathcal{C}\log_2(\lambda_S)$ for some universal constant $\mathcal{C}$ (large).

Assume that $R > 0$ is the distance between b and the set S. Let $\varepsilon_i, \Delta_i, i = 0, 1, 2\ldots$ and $R = r_0 < r_1 < r_2 < \ldots$ be positive scalars (their values will be defined later). For each $k = 1, 2, 3, \ldots$ we define an annulus

$$X_k = S \cap B(b, r_k) \setminus B(b, r_{k-1}).$$

Since $X_k \subseteq S \cap B(b, r_k)$, by Lemma 5.3.2 we can find a point set $S_k \subseteq S$ of size $|S_k| \leq \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k}) \rceil}$ such that

$$X_k \subseteq \bigcup_{s \in S_k} B(s, \varepsilon_k).$$

Hence, for any $x \in X_k$, there is $s \in S_k$ such that $\|x - s\| < \varepsilon_k$. Moreover, by the triangle inequality, any such s satisfies $r_{k-1} - \varepsilon_k < \|s - b\| < r_k + \varepsilon_k$ (since x is inside the annuli X_k). So without loss of generality we can assume that

$$S_k \subseteq B(b, r_k + \varepsilon_k) \setminus B(b, r_{k-1} - \varepsilon_k).$$

Using the union bound, we have:

$$\begin{aligned} \mathsf{Prob}\big(\exists x \in S \text{ s.t } T(x) = T(b)\big) &= \mathsf{Prob}\big(\exists x \in \bigcup_{k=1}^{\infty} X_k \text{ s.t } T(x) = T(b)\big) \\ &\leq \sum_{k=1}^{\infty} \mathsf{Prob}\big(\exists x \in X_k \text{ s.t } T(x) = T(b)\big). \end{aligned}$$

Now, we will try to estimate the individual probabilities inside this sum.

For each $k \geq 1$, we denote by $\mathcal{E}_k$ the event that:

$$\exists s \in S_k,\ \exists x \in X_k \cap B(s, \varepsilon_k) \text{ s.t. } \|Ts - Tx\| > \Delta_k.$$

Then we have

$$\mathsf{Prob}\big(\exists x \in X_k \text{ s.t } T(x) = T(b)\big) \leq \mathsf{Prob}\big((\exists x \in X_k \text{ s.t } T(x) = T(b)) \wedge \mathcal{E}_k^c\big) + \mathsf{Prob}(\mathcal{E}_k) \quad (5.12)$$

For the second term in (5.12), by the union bound, we have

$$\mathsf{Prob}(\mathcal{E}_k) \leq \sum_{s \in S_k} \mathsf{Prob}\big(\exists x \in X_k \cap B(s, \varepsilon_k) \text{ s.t. } \|Ts - Tx\| > \Delta_k\big).$$

If we denote by $X_k^s = \{x - s |\ x \in X_k\}$, then the RHS can be written as

$$\begin{aligned}
& \sum_{s\in S_k} \mathsf{Prob}\big(\exists (x-s) \in X_k^s \cap B(0,\varepsilon_k) \text{ s.t. } \|Ts - Tx\| > \Delta_k\big) \\
= & \sum_{s\in S_k} \mathsf{Prob}\big(\exists u \in X_k^s \cap B(0,\varepsilon_k) \text{ s.t. } \|Tu\| > \Delta_k\big) \\
\leq & \sum_{s\in S_k} e^{-c_1 d(\frac{\Delta_k}{\varepsilon_k})^2} \qquad \text{(for the universal constant } c_1 \text{ in Lemma 5.3.3)} \\
\leq & \ \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil}\ e^{-c_1 d(\frac{\Delta_k}{\varepsilon_k})^2}.
\end{aligned}$$

(Note that here we must choose $\Delta_k \geq 6\varepsilon_k$ in order to apply the Lemma 5.3.3).

For the first term in (5.12), we have

$$\begin{aligned}
& \mathsf{Prob}\big((\exists x \in X_k \text{ s.t } T(x) = T(b)) \wedge \mathcal{E}_k^c\big) \\
\leq & \ \mathsf{Prob}\big(\exists x \in X_k, s \in S_k \cap B(x,\varepsilon_k) \text{ s.t } T(x) = T(b) \wedge \|T(s) - T(x)\| \leq \Delta_k\big) \\
\leq & \ \mathsf{Prob}\big(\exists s \in S_k \text{ s.t } \|T(s) - T(b)\| < \Delta_k\big) \\
\leq & \ \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil}\ \mathsf{Prob}\big(\|T(z)\| < \frac{\Delta_k}{r_{k-1} - \varepsilon_k}\big) \quad \text{for some unit vector } z \\
\leq & \ \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil}\ (\frac{\Delta_k}{r_{k-1} - \varepsilon_k})^d \quad \text{(by inequality 5.8).}
\end{aligned}$$

Putting all the estimations we have obtained together, we have:

$$\mathsf{Prob}\big(\exists x \in S \text{ s.t } T(x) = T(b)\big) \leq \sum_{k=1}^{\infty} \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil} \left(e^{-c_1 d(\frac{\Delta_k}{\varepsilon_k})^2} + (\frac{\Delta_k}{r_{k-1} - \varepsilon_k})^d \right). \tag{5.13}$$

Next, we will show that there are choices of $\varepsilon_k, \Delta_k, r_k$ such that the RHS of (5.13) can be as small as needed.

Choices of $\varepsilon_k, \Delta_k, r_k$: For some $N > 1$ large, we will choose $\varepsilon_k, \Delta_k, r_k$ as follows:

1. $\varepsilon_k = \varepsilon$, for some $\varepsilon > 0$.
2. $\Delta_k = Nk\varepsilon$
3. $r_k = (N^2(k+1)^2 + 1)\varepsilon$

Now the RHS of (5.13) can be rewritten as follows

$$\begin{aligned}
& \sum_{k=1}^{\infty} \lambda_S^{\lceil \log_2(N^2(k+1)^2+1)\rceil} \left(e^{-c_1 d(Nk)^2} + (\frac{1}{Nk})^d \right) \\
\leq & \sum_{k=1}^{\infty} \lambda_S^{3\log_2(Nk)} \left(e^{-c_1 d(Nk)^2} + (\frac{1}{Nk})^d \right) \\
\leq & \sum_{k=1}^{\infty} 2^{3\,\mathsf{ddim}(S)\log_2(Nk)} \left(e^{-c_1 d(Nk)^2} + (\frac{1}{Nk})^d \right).
\end{aligned} \tag{5.14}$$

Note that $2^{3\,\mathsf{ddim}(S)\log_2(Nk)}$ does not increase fast enough compared to the decreasing speeds of both $e^{-c_1 d(Nk)^2}$ and $(\frac{1}{Nk})^d$ when $d \geq \mathcal{C}\mathsf{ddim}(S)$ with $\mathcal{C}$ large enough (and also independent of N). Therefore, there are universal constants $c_2, c_3 > 0$ such that the value of (5.14) is less than or equal to

$$\sum_{k=1}^{\infty} e^{-c_2(Nk)^2} + \sum_{k=1}^{\infty} (\frac{1}{Nk})^{c_3 d} \tag{5.15}$$

Both the two infinite sums tend to 0 when N tends to ∞. This means that

$$\mathsf{Prob}\big(\exists x \in S \text{ s.t } T(x) = T(b)\big) = 0,$$

which proves our theorem. □

Our next result in this section is an extension of Thm. 5.3.4 to the threshold case.

Theorem 5.3.5. *Let $b \notin S$ where $S \subseteq \mathbb{R}^m$ is a closed doubling set, $T : \mathbb{R}^m \to \mathbb{R}^d$ be a Gaussian random projection, and $r = \min_{x \in S} \|b - x\|$. Let $\kappa < 1$ be some fixed constant. Then for all $0 < \delta < 1$ and $0 < \tau < \kappa r$, we have*

$$\mathsf{Prob}(\mathsf{dist}(T(p), T(S)) > \tau) > 1 - \delta$$

if the projected dimension $d = \Omega(\frac{\log(\frac{\lambda_S}{\delta})}{1-\kappa})$.

Proof. Let $d \geq \mathcal{C}(\frac{\log(\frac{\lambda_S}{\delta})}{1-\kappa})$ for some universal constant $\mathcal{C}$ (large). As before, let $\varepsilon_i, \Delta_i, i = 0, 1, 2 \ldots$ and $r = r_0 < r_1 < r_2 < \ldots$ be positive scalars whose values will be decided later. The annuli X_k and point sets S_k are also defined similarly. Using the union bound, now we have:

$$\begin{aligned} \mathsf{Prob}\big(\exists x \in S \text{ s.t } \|T(x) - T(b)\| < \tau\big) &= \mathsf{Prob}\big(\exists x \in \bigcup_{k=1}^{\infty} X_k \text{ s.t } \|T(x) - T(b)\| < \tau\big) \\ &\leq \sum_{k=1}^{\infty} \mathsf{Prob}\big(\exists x \in X_k \text{ s.t } \|T(x) - T(b)\| < \tau\big). \end{aligned}$$

Now, we will try to estimate the individual probabilities inside this sum.

For each $k \geq 1$, we denote by $\mathcal{E}_k$ the event that:

$$\exists s \in S_k,\ \exists x \in X_k \cap B(s, \varepsilon_k) \text{ s.t. } \|Ts - Tx\| > \Delta_k. \tag{5.16}$$

Then we have

$$\begin{aligned} & \mathsf{Prob}\big(\exists x \in X_k \text{ s.t } \|T(x) - T(b)\| < \tau\big) \\ \leq\ & \mathsf{Prob}\big((\exists x \in X_k \text{ s.t } \|T(x) - T(b)\| < \tau) \wedge \mathcal{E}_k^c\big) + \mathsf{Prob}(\mathcal{E}_k) \end{aligned} \tag{5.17}$$

For the second term in (5.17), from the previous proof, we already had:

$$\mathsf{Prob}(\mathcal{E}_k) \le \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil} e^{-c_1 d(\frac{\Delta_k}{\varepsilon_k})^2}. \tag{5.18}$$

(Note that here we must choose $\Delta_k \ge 6\varepsilon_k$ in order to apply the Lemma 5.3.3).

Now, for the first term in (5.17), we have

$$\begin{aligned}
&\mathsf{Prob}\big((\exists x \in X_k \text{ s.t } \|T(x) - T(b)\| < \tau) \wedge \mathcal{E}_k^c\big) \\
\le\ &\mathsf{Prob}\big(\exists x \in X_k, s \in S_k \cap B(x, \varepsilon_k) \text{ s.t } \|T(x) - T(b)\| < \tau \wedge \|T(s) - T(x)\| \le \Delta_k\big) \\
\le\ &\mathsf{Prob}\big(\exists s \in S_k \text{ s.t } \|T(s) - T(b)\| < \Delta_k + \tau\big) \qquad \text{(by triangle inequality)} \\
\le\ &\begin{cases} \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil} \mathsf{Prob}\big(\|T(z)\| < \frac{\Delta_k+\tau}{r_{k-1}-\varepsilon_k}\big) & \text{for some unit vector } z \quad \text{if } k \ge 2 \\ \lambda_S^{\lceil \log_2(\frac{r_1}{\varepsilon_1})\rceil} \mathsf{Prob}\big(\|T(z)\| < \frac{\Delta_1+\tau}{r}\big) & \text{for some unit vector } z \quad \text{if } k = 1 \end{cases} \\
\le\ &\begin{cases} \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil} (\frac{\Delta_k+\tau}{r_{k-1}-\varepsilon_k})^d & \text{if } k \ge 2 \\ \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil} (\frac{\Delta_1+\tau}{r})^d & \text{if } k = 1. \end{cases}
\end{aligned}$$

Putting all the estimations we have obtained together, we have:

$$\begin{aligned}
&\mathsf{Prob}\big(\exists x \in S \text{ s.t } \|T(x) - T(b)\| < \tau\big) \\
\le\ &\left(\sum_{k=1}^{\infty} \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil} e^{-c_1 d(\frac{\Delta_k}{\varepsilon_k})^2} + \sum_{k=2}^{\infty} \lambda_S^{\lceil \log_2(\frac{r_k}{\varepsilon_k})\rceil} (\frac{\Delta_k + \tau}{r_{k-1} - \varepsilon_k})^d\right) + \lambda_S^{\lceil \log_2(\frac{r_1}{\varepsilon_1})\rceil} (\frac{\Delta_1 + \tau}{r})^d
\end{aligned} \tag{5.19}$$

.

Here we separate one term out, and we will prove that the remaining expression can be made as small as wanted by certain choices of parameters.

Choices of $\varepsilon_k, \Delta_k, r_k$: Let $N > 0$ be the number such that $(\frac{7}{N}+1) = \frac{r}{\tau}$. From the assumption, we have $N < \frac{7\kappa}{1-\kappa} < \frac{7}{1-\kappa}$. We will choose $\varepsilon_k, \Delta_k, r_k$ as follows:

1. $\varepsilon_k = \varepsilon = \frac{\tau}{N}$,
2. $\Delta_k = 6\sqrt{k}\varepsilon$,
3. $r_k = (6k + 7)\varepsilon + \sqrt{k+1}\,\tau$.

(Our purpose is to choose the parameters so that $\frac{\Delta_k+\tau}{r_{k-1}-\varepsilon_k} = \frac{1}{\sqrt{k}}$ and $\Delta_k \ge 6\varepsilon_k$).

From this choice, it is obvious $r_0 = r$. Now the RHS of (5.19) can be rewritten as follows

$$\begin{aligned}
&\left(\sum_{k=1}^{\infty} \lambda_S^{\lceil \log_2(6k+7+N\sqrt{k+1})\rceil} e^{-c_1 d(36k^2)} + \sum_{k=2}^{\infty} \lambda_S^{\lceil \log_2(6k+7+N\sqrt{k+1})\rceil} (\frac{1}{\sqrt{k}})^d\right) + \lambda_S^{\lceil \log_2(13+N\sqrt{2})\rceil} ((\frac{6}{N}+1)\frac{\tau}{r})^d. \\
\le\ &\left(\sum_{k=1}^{\infty} \lambda_S^{c_3 \log_2(N(k+1))} e^{-c_1 d(36k^2)} + \sum_{k=2}^{\infty} \lambda_S^{c_3 \log_2(N(k+1))} (\frac{1}{\sqrt{k}})^d\right) + \lambda_S^{c_2} ((\frac{6}{N}+1)\frac{\tau}{r})^d
\end{aligned} \tag{5.20}$$

for some universal constants c_1, c_2, c_3.

It is easy to show that the expression is the big bracket is bounded above by $e^{-c_4 d}$ as long as $d \geq \mathcal{C} \log_2(\lambda_S) \log(\frac{7}{1-\kappa}) > \mathcal{C} \log_2(\lambda_S) \log(N)$ (for some large constants $c_4, \mathcal{C}$). Moreover,

$$e^{-c_4 d} \leq \frac{\delta}{2} \quad \text{if and only if} \quad d \geq \frac{1}{c_4} \log(\frac{2}{\delta}) \tag{5.21}$$

and $\lambda_S^{c_2} \ ((\frac{6}{N} + 1)\frac{\tau}{r})^d \leq \frac{\delta}{2}$ if and only if

$$d \geq \frac{c_2 \log(\lambda_S) + \log(\frac{2}{\delta})}{\log((\frac{N}{6+N})\frac{r}{\tau})}. \tag{5.22}$$

However, $\log((\frac{N}{6+N})\frac{r}{\tau}) = \log(\frac{7r}{6r+\tau}) \geq \log(\frac{7}{6+\kappa}) \geq \log(1 + \frac{1-\kappa}{7}) \geq \frac{1-\kappa}{7}$, from the Taylor series of the logarithm function. Therefore, (5.22) holds if we select

$$d \geq \mathcal{C} \frac{\log(\frac{\lambda_S}{\delta})}{1-\kappa} \tag{5.23}$$

for some universal constants $\mathcal{C}$.

The proof follows immediately from (5.21) and (5.23) by apply the union bound. □

One of the interesting consequence of Theorem 5.3.5 is the following application to the Approximate Nearest Neighbour problem.

Corollary 5.3.6. *For $X \subseteq \mathbb{R}^d$, $\varepsilon \in (0, 1/2)$ and $\delta \in (0, 1/2)$, there exists*

$$d = \max\left\{ O\left(\frac{\log(\frac{1}{\delta})}{\varepsilon^2}\right),\ O\left(\frac{\log(\frac{\lambda_S}{\delta})}{\varepsilon}\right) \right\}$$

such that for every $x_0 \in X$, with probability at least $1 - \delta$,

1. $\mathsf{dist}\ (Tx_0, T(X \setminus \{x_0\})) \leq (1+\varepsilon)\mathsf{dist}\ (x_0, X \setminus \{x_0\})$,
2. *Every $x \in X$ with $\|x_0 - x\| \geq (1+2\varepsilon)\mathsf{dist}\ (x_0, X \setminus \{x0\})$ satisfies*

$$\|Tx_0 - Tx\| > (1+\varepsilon)\mathsf{dist}\ (x_0, X \setminus \{x0\}).$$

Note that, this result improves the bound provided by Indyk-Naor in ([12]). In that paper, the authors give the bound for the projected dimension to be

$$d = O\left(\frac{\log(2/\varepsilon)}{\varepsilon^2} \log(\frac{1}{\delta}) \log(\lambda_S)\right),$$

which is significantly larger than our bound.

Proof. Similar as the proof of Theorem 4.1 in ([12]), we have: for $d \geq \mathcal{C}\frac{\log(\frac{1}{\delta})}{\varepsilon^2}$ with some large constant $\mathcal{C}$:

$$\text{Prob}\left[\mathsf{dist}\left(Tx_0, T(X \setminus \{x_0\})\right) \leq (1+\varepsilon)\mathsf{dist}\left(x_0, X \setminus \{x_0\}\right)\right] < \frac{\delta}{2}. \tag{5.24}$$

Now, assume that $\mathsf{dist}\left(x_0, X \setminus \{x0\}\right) = 1$. Let set $b = x_0$ and $S = \{x \in X \ : \ \|x - x_0\| \geq 1+2\varepsilon\}$ and $\tau = 1+\varepsilon$ as in Theorem 5.3.5. We then have $r := \min_{x \in S} \|x - b\| \geq 1 + 2\varepsilon$, which implies $\frac{\tau}{r} \leq \frac{1+\varepsilon}{1+2\varepsilon} = 1 - \frac{\varepsilon}{1+2\varepsilon} < 1 - \frac{\varepsilon}{2}$. Therefore, we can choose $\kappa = 1 - \varepsilon/2$. Applying Theorem 5.3.5, we have:

$$\mathsf{Prob}(\mathsf{dist}(T(b), T(S)) \leq \tau) \leq \frac{\delta}{2} \tag{5.25}$$

if the projected dimension $d = \Omega(\frac{\log(\frac{\lambda_S}{\delta})}{1-\kappa}) = \Omega(\frac{\log(\frac{\lambda_S}{\delta})}{\varepsilon})$.

From (5.24) and (5.25), we conclude that the two required conditions hold for some

$$d = \max\left\{O(\frac{\log(\frac{1}{\delta})}{\varepsilon^2}),\ O(\frac{\log(\frac{\lambda_S}{\delta})}{\varepsilon})\right\},$$

as claimed. □

Chapter 6

Random projections for trust-region subproblems

6.1 Derivative-free optimization and trust-region methods

Derivative-free optimization (DFO) is a branch of optimization that has attracted a lot of attentions recently. In the general form, a DFO problem is defined as follows:

$$\min \{f(x) \mid x \in \mathcal{D}\},$$

in which $\mathcal{D}$ is a subset of $\mathbb{R}^n$ and $f(.)$ is a continuous function such that no derivative information about it is available. In many cases, we have to treat the objective function as a black-box. It means that we can only understand $f(.)$ by evaluating it at a limited number of input points.

Because of the lack of derivative information, traditional gradient-based methods cannot be applied. Moreover, when the objective function is expensive, meta-heuristics such as evolutionary algorithms, simulated annealing or ant colony optimization (ACO) are not desirable, since they often require a very large number of function evaluations. Recently, trust-region (TR) method stands out as one of the most efficient methods for solving DFO. TR methods involve the construction of surrogate models to approximate the true function (locally) in small subsets of $\mathcal{D}$ and relying on those models to search for optimal solutions. These local subsets are called trust-regions and often chosen as closed balls with respect to certain norms. There are several ways to update the new data points, but the most common way is to find them as minima of the current model over the current trust-region. Formally, we need to

solve the following problems, which are often called trust-region sub-problems:

$$\min \{m(x) \mid x \in B(c, r) \cap \mathcal{D}\}.$$

Here the balls $B(c, r)$ and the models $m(.)$ are updated at every iteration. The newly found solutions of these problems will then be evaluated under the objective function $f(.)$. Based on their values, we can adaptively adjust the model as well as the trust region, i.e. either to expand, contract, or switch their centers ...

The models $m(.)$ are often chosen to be simple so that the TR subproblems are easy to solve. The most common choices for $m(.)$ are linear and quadratic functions. However, when the instances are large, solving TR subproblems becomes difficult and sometimes they are the bottle-neck in applying TR methods for large-scale problems. For example, it is known that quadratic programming is NP-hard, even with only one negative eigenvalue.

In this chapter, we propose to use random projections to speed up the solving of high-dimensional TR subproblems. We assume that linear and quadratic functions are used as TR models and $\| \, . \, \|_2$ is used for norm. Moreover, we assume that $\mathcal{D}$ is a polyhedron defined by explicit linear inequality constraints. When scaling properly, a TR subproblem can be rewritten as

$$\min \{x^\top Q x + c^\top x \mid Ax \leq b, \ \|x\|_2 \leq 1\}, \tag{6.1}$$

in which $Q \in \mathbb{R}^{n \times n}, \ A \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m$.

Now, let $P \in \mathbb{R}^{d \times n}$ be a random projection with i.i.d. $\mathcal{N}(0, 1)$ entries, we can "project" x to Px and study the following projected problem

$$\min \{x^\top (P^\top P Q P^\top P) x + c^\top P^\top P x \mid AP^\top P x \leq b, \ \|Px\|_2 \leq 1\}.$$

By setting $u = Px, \ \bar{c} = Pc, \ \overline{A} = AP^\top$, we can rewrite it as

$$\min \{u^\top (P Q P^\top) u + \bar{c}^\top u \mid \overline{A} u \leq b, \ \|u\|_2 \leq 1\}. \tag{6.2}$$

This problem is of much lower dimension than the original one, so it is expected to be much easier. However, as we will show later, with high probability, it still provides us a good approximate solution. This result is interesting, especially for DFO and black-box problems. In these cases, it is unwise to spend too much time on solving TR subproblems and we are often happy with approximate solutions. Moreover, since the surrogate models might not even fit the true objective function, we are more or less tolerative to the small probability of failure.

6.2 Random projections for linear and quadratic models

In this section, we will explain the motivations for the study of the projected problem (6.2). We start with the following simple lemma, which says that linear and quadratic models can be approximated well under random projections:

6.2.1 Approximation results

Lemma 6.2.1. *Let $P : \mathbb{R}^n \to \mathbb{R}^d$ be a random projection and $0 < \varepsilon < 1$. Then there is a constant C such that*

(i) For any $x, y \in \mathbb{R}^n$:

$$\langle Px, Py\rangle = \langle x, y\rangle \pm \varepsilon\|x\|.\|y\|$$

with probability at least $1 - 4e^{-C\varepsilon^2 d}$.

(ii) For any $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times n}$ whose rows are unit vectors:

$$AP^TPx = Ax \pm \varepsilon\|x\| \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix}$$

with probability at least $1 - 4me^{-C\varepsilon^2 d}$.

(iii) For any two vectors $x, y \in \mathbb{R}^n$ and a square matrix $Q \in \mathbb{R}^{n\times n}$, then with probability at least $1 - 8ke^{-C\varepsilon^2 d}$, we have:

$$x^TP^TPQP^TPy = x^TQy \pm 3\varepsilon\|x\|.\|y\|.\|Q\|_*,$$

in which $\|Q\|_*$ is the nuclear norm of Q and k is the rank of Q.

Proof.
(i) This property has already been proved in Lemma 2.2.2, Chapter 2.

(ii) Let $A_1, \dots, A_m$ be (unit) row vectors of A. Then

$$AP^\top Px - Ax = \begin{pmatrix} A_1^\top P^\top Px - A_1^\top x \\ \dots \\ A_m^\top P^\top Px - A_m^\top x \end{pmatrix} = \begin{pmatrix} \langle PA_1, Px\rangle - \langle A_1, x\rangle \\ \dots \\ \langle PA_m, Px\rangle - \langle A_m, x\rangle \end{pmatrix}.$$

The claim is then followed by applying Part (i) and the union bound.

(iii) Let $Q = U\Sigma V^T$ be the Singular Value Decomposition of Q. Here U, V are $(n \times k)$-real matrices with orthogonal unit column vectors $u_1, \ldots, u_k$ and $v_1, \ldots, v_k$, respectively and $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_k)$ is a diagonal matrix with positive entries. Denote by $\mathbf{1}_k = (1, \ldots, 1)^\top$ the k-dimensional column vector of all 1 entries. Then

$$\begin{aligned} x^T P^T PQP^T Py &= (U^\top P^\top Px)^\top \Sigma (V^\top P^\top Py) \\ &= (U^\top x \pm \varepsilon \|x\| \mathbf{1}_k)^\top \, \Sigma \, (V^\top y \pm \varepsilon \|y\| \mathbf{1}_k) \end{aligned}$$

with probability at least $1 - 8ke^{-C\varepsilon^2 d}$ (by applying part (ii) and the union bound). Moreover

$$\begin{aligned} & (U^\top x \pm \varepsilon \|x\| \mathbf{1}_k)^\top \, \Sigma \, (V^\top y \pm \varepsilon \|y\| \mathbf{1}_k) \\ = \; & x^T Qy \pm \varepsilon \|x\| . (\mathbf{1}_k^\top \Sigma V^\top y) \pm \varepsilon \|y\| . (x^\top U \Sigma \mathbf{1}_k) \pm \varepsilon^2 \|x\| . \|y\| . \sum_{i=1}^k \sigma_i \\ = \; & x^T Qy \pm \varepsilon (\sigma_1, \ldots, \sigma_k) \big(\|x\| V^\top y \pm \|y\| U^\top x \big) \pm \varepsilon^2 \|x\| . \|y\| . \sum_{i=1}^k \sigma_i . \end{aligned}$$

Therefore,

$$\begin{aligned} |x^T P^T PQP^T Py - x^T Qy| \leq \; & \|x\| . \|y\| . (2\varepsilon \sqrt{\sum_{i=1}^k \sigma_i^2} + \varepsilon^2 \sum_{i=1}^k \sigma_i) \\ \leq \; & 3\varepsilon \|x\| . \|y\| . \sum_{i=1}^k \sigma_i = 3\varepsilon \|x\| . \|y\| . \|Q\|_* \end{aligned}$$

with probability at least $1 - 8ke^{-C\varepsilon^2 d}$. □

It is known that singular values of random matrices often concentrate around their expectations. In the case when the random matrix is sampled from Gaussian ensemble, this phenomenon is well-understood due to many current research efforts. The following lemma, which is proved in [28], uses this phenomenon to show that when $P \in \mathbb{R}^{d \times n}$ is a Gaussian random matrix (with the number of row significantly smaller than the number of columns), then $PP^\top$ is very close to the identity matrix.

Lemma 6.2.2. *Let $P \in \mathbb{R}^{d \times n}$ be a random matrix in which each entry is an i.i.d $\mathcal{N}(0, \frac{1}{\sqrt{n}})$ random variable. Then for any $\delta > 0$ and $0 < \varepsilon < \frac{1}{2}$, with probability at least $1 - \delta$, we have:*

$$\|PP^\top - I\|_2 \leq \varepsilon$$

provided that

$$n \geq \frac{(d+1)\log(\frac{2d}{\delta})}{c\varepsilon^2},$$

where $\| \, . \, \|_2$ is the spectral norm of matrix and $c > \frac{1}{4}$ is some universal constant.

This lemma also tells us that when we go from low to high dimensions, with high probability we can ensure that the norms of all the points endure small distortions. Indeed, for any vector $u \in \mathbb{R}^d$, then

$$\|P^\top u\|^2 - \|u\|^2 = \langle P^\top u, P^\top u\rangle - \langle u, u\rangle = \langle (PP^\top - I)u, u\rangle = \pm\varepsilon\|u\|^2,$$

due to the Cauchy–Schwarz inequality. Moreover, it implies that $\|P\|_2 \leq (1+\varepsilon)$ with probability at least $1-\delta$.

6.2.2 Trust-region subproblems with linear models

We will first work with a simple case, when the surrogate models using in TR methods are linear and defined as follows:

$$\min\{c^\top x \mid Ax \leq b,\ \|x\|_2 \leq 1, x \in \mathbb{R}^n\}. \tag{6.3}$$

We will establish the relations between problem (6.3) and its corresponding projected problems:

$$\min\{(Pc)^\top u \mid AP^\top u \leq b,\ \|u\|_2 \leq 1-\varepsilon, u \in \mathbb{R}^d\} \tag{P_ε^-}$$

We first obtain the following feasibility result:

Theorem 6.2.3. *Let $P \in \mathbb{R}^{d\times n}$ be a random matrix in which each entry is an i.i.d $\mathcal{N}(0, \frac{1}{\sqrt{n}})$ random variable. Let $\delta \in (0,1)$. Assume further that*

$$n \geq \frac{(d+1)\log(\frac{2d}{\delta})}{c\varepsilon^2},$$

for some universal constant $c > \frac{1}{4}$. Then with probability at least $1-\delta$, for any feasible solution u of the projected problem (P_ε^-), $P^\top u$ is also feasible for the original problem (6.3).

We should notice a universal property in this theorem, in which with a fixed probability, the feasibility holds for "all" (instead of a specific) vectors u.

Proof. Let u be any feasible solution for the projected problem (P_ε^-) and take $\hat{x} = P^\top u$. Then we have $A\hat{x} = AP^\top u \leq b$ and

$$\|\hat{x}\|^2 = \|P^\top u\|^2 = u^\top P^\top P u = u^\top u + u^\top(P^\top P - I)u \leq (1+\varepsilon)\|u\|^2,$$

with probability at least $1-\delta$ (by Lemma 6.2.2). Since $\|u\| \leq 1-\varepsilon$, we have

$$\|\hat{x}\| \leq (1+\varepsilon)(1-\varepsilon) < 1,$$

which proves the theorem. □

In order to estimate the quality of the objective values, we define another projected problem, which can be considered as a relaxation for the previous one:

$$\min\{(Pc)^\top u \mid AP^\top u \leq b + \varepsilon,\ \|u\|_2 \leq 1 + \varepsilon, u \in \mathbb{R}^d\} \tag{P_ε^+}$$

Intuitively, these two projected problems are very close to each other when ε is small enough (under some additional assumptions). Moreover, for practical use, we need to assume that they are both feasible. Let u_ε^- and u_ε^+ be optimal solutions for these two problems, respectively. Denote by $x_\varepsilon^- = P^\top u_\varepsilon^-$ and $x_\varepsilon^+ = P^\top u_\varepsilon^+$. Let x^* be an optimal solution for the original problem (6.1). We will try to bound $c^\top x^*$ between $c^\top x_\varepsilon^-$ and $c^\top x_\varepsilon^+$, the two values that are expected to be approximately close to each other.

Theorem 6.2.4. *Let $P \in \mathbb{R}^{d\times n}$ be a random matrix in which each entry is an i.i.d $\mathcal{N}(0, \frac{1}{\sqrt{n}})$ random variable. Let $\delta \in (0,1)$. Assume further that*

$$n \geq \frac{(d+1)\log(\frac{2d}{\delta})}{c\varepsilon^2},$$

for some universal constant $c > \frac{1}{4}$ and

$$d \geq \mathcal{C}\frac{\log(m/\delta)}{\varepsilon^2},$$

for some universal constant $\mathcal{C} > 1$.
Let x^* be an optimal solution for the original problem (6.1). Then
(i) With probability at least $1-\delta$, the solution x_ε^- is feasible for the original problem (6.1).
(ii) With probability at least $1-\delta$, we have:

$$c^T x_\varepsilon^- \geq c^T x^* \geq c^T x_\varepsilon^+ - \varepsilon\|c\|.$$

Proof. (i) From the previous theorem, with probability at least $1-\delta$, for any feasible point u of the projected problem (P_ε^-), $P^\top u$ is also feasible for the original problem (6.3). Therefore, it must hold also for x_ε^-.

(ii) From part (i), with probability at least $1-\delta$, x_ε^- is feasible for the original problem (6.1). Therefore, we have

$$c^T x_\varepsilon^- \geq c^\top x^*,$$

with probability at least $1-\delta$.

Moreover, due to Lemma 6.2.1, with probability at least $1 - 4e^{-\mathcal{C}\varepsilon^2 d}$, we have

$$c^\top x^* \geq c^\top P^\top P x^* - \varepsilon\|c\|.\|x^*\| \geq c^\top P^\top P x^* - \varepsilon\|c\|,$$

since $\|x^*\| \leq 1$. On the other hand, let $\hat{u} := Px^*$, due to Lemma 6.2.1, we have

$$AP^\top \hat{u} = AP^\top Px^* \leq Ax^* + \varepsilon \|x^*\| \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix} \leq Ax^* + \varepsilon \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix} \leq b + \varepsilon,$$

with probability at least $1 - 4me^{-C\varepsilon^2 d}$, and

$$\|\hat{u}\| = \|Px^*\| \leq (1+\varepsilon)\|x^*\| \leq (1+\varepsilon),$$

with probability at leats $1 - 2e^{-C\varepsilon^2 d}$ (this is the norm preservation property of random projections). Therefore, $\hat{u}$ is a feasible solution for the problem (P_ε^+) with probability at least $1 - (4m+2)e^{-C\varepsilon^2 d}$. Due to the optimality of u_ε^+ for the problem (P_ε^+), it follows that

$$c^\top x^* \geq c^\top P^\top Px^* - \varepsilon\|c\| = c^\top P^\top \hat{u} - \varepsilon\|c\| \geq c^\top P^\top u_\varepsilon^+ - \varepsilon\|c\| = c^\top x_\varepsilon^+ - \varepsilon\|c\|,$$

with probability at least $1-(4m+6)e^{-C\varepsilon^2 d}$, which is at least $1-\delta$ for some universal constant C. □

We have established that $c^\top x^*$ is sandwiched between $c^T x_\varepsilon^-$ and $c^T x_\varepsilon^+$. Now we will compare these two values. For simplicity, we assume that the feasible set

$$S^* = \{x \in \mathbb{R}^n \mid Ax \leq b,\ \|x\|_2 \leq 1\}$$

is of full dimension. We associate with each set P a positive number $\text{FULL}(P) > 0$, which is considered as a fullness measure of P and is defined as the maximum radius of any closed ball contained in P. Now, assume that $\text{FULL}(S^*) = r^* > 0$.

The following lemma characterizes the fullness of S_ε^+ with respect to r^*, in which

$$S_\varepsilon^+ := \{u \in \mathbb{R}^d \mid AP^\top u \leq b + \varepsilon,\ \|u\|_2 \leq 1 + \varepsilon\},$$

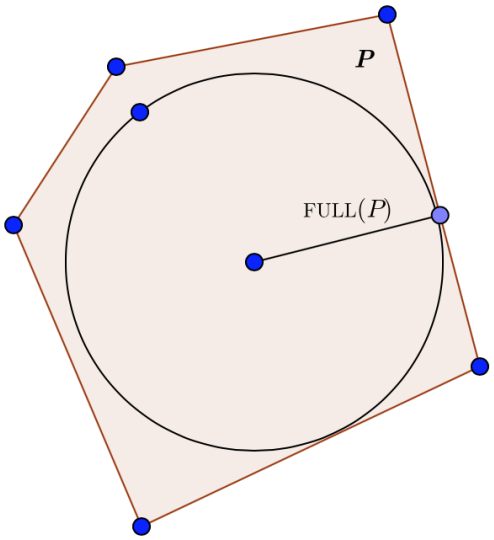
that is, the feasible set of the problem (P_ε^+).

Lemma 6.2.5. *Let S^* be of full dimension with $\text{FULL}(S^*) = r^*$. Then with probability at least $1 - 3\delta$, S_ε^+ is also full dimensional with the fullness measure:*

$$\text{FULL}(S_\varepsilon^+) \geq (1-\varepsilon)r^*.$$

In the proof of this lemma, we will extensively use the fact that, by Cauchy-Schwartz, for any row vector $a \in \mathbb{R}^n$

$$\sup_{\|u\| \leq r} au = r\|a\|.$$



Proof. For any $i \in [1, ..., n]$ let A_i denotes the ith row of A. Let $B(x_0, r^*)$ be a closed ball contained in S^*. Then for any $x \in \mathbb{R}^n$ with $\|x\| \leq r^*$, we have $A(x_0 + x) = Ax_0 + Ax \leq b$, which implies that for any $i \in [1, ..., n]$,

$$b_i \geq (Ax_0)_i + \sup_{\|x\| \leq r^*} A_i x = (Ax_0)_i + r^* \|A_i\| = (Ax_0)_i + r^*, \tag{6.4}$$

hence

$$b \geq Ax_0 + r^*$$

By Lemma 6.2.1, with probability at least $1 - \delta$, we have

$$AP^\top Px_0 \leq Ax_0 + \varepsilon \leq b + \varepsilon - r^*.$$

Let $u \in \mathbb{R}^n$ with $\|u\| \leq (1 - \varepsilon)r^*$, for any $i \in [1, ..., n]$, we have:

$$(AP^\top(Px_0 + u))_i \leq b_i + \varepsilon - r^* + (AP^\top)_i u = b_i + \varepsilon - r^* + A_i P^\top u$$

where $(AP^\top)_i$ denotes the ith row of $AP^\top$. Hence, by Cauchy-Schwartz,

$$(AP^\top(Px_0 + u))_i \leq b_i + \varepsilon - r^* + (1 - \varepsilon)r^* \|A_i P^\top\|$$

Using Eq. (**??**), we can prove that with probability $1 - 2me^{-\mathcal{C}\varepsilon^2 d} \geq 1 - \delta$, we have that for all $i \in [1, ..., m]$, $\|A_i P^\top\| \leq (1 + \varepsilon)\|A_i\| = (1 + \varepsilon)$. Hence

$$AP^\top(Px_0 + u) \leq b_i + \varepsilon - r^* + (1 - \varepsilon)r^*(1 + \varepsilon) \leq b + \varepsilon$$

therefore, with probability $1 - 2\delta$ the closed ball B^* centered at Px_0 with radius $(1 - \varepsilon)r^*$ is contained in $\{u : AP^\top u \leq b + \varepsilon\}$.

Moreover, since $B(x_0, r^*)$ is contained in S^*, which is the subset of the unit ball, then $\|x_0\| \leq 1 - r^*$.

With probability at least $1 - \delta$, for all vectors u in $\textsc{ball}(Px_0, (1-\varepsilon)r^*)$, we have

$$\|u\| \leq \|Px_0\| + (1-\varepsilon)r^* \leq (1+\varepsilon)\|x_0\| + (1-\varepsilon)r^* \leq (1+\varepsilon)(1-r^*) + (1-\varepsilon)r^* \leq 1 + \varepsilon.$$

Therefore, by the definition of S_ε^+ we have

$$\textsc{ball}\big(Px_0, (1-\varepsilon)r^*\big) \subseteq S_\varepsilon^+,$$

which implies that the fullness of S_ε^+ is at least $(1-\varepsilon)r^*$, with probability at least $1 - 3\delta$. $\square$

Now, assume that $B(u_0, r_0)$ is a closed ball with maximum radius that is contained in S_ε^+.

In order to establish the relation between u_ε^+ and u_ε^-, our idea is to move u_ε^+ a bit closer to u_0 (defined in the above lemma), so that the new point is contained in S_ε^-. Therefore, its objective value will be at least that of u_ε^-, but quite close to the objective value of u_ε^+.

We define $\hat{u} := (1-\lambda)u_\varepsilon^+ + \lambda u_0$ for some $\lambda \in (0,1)$ specified later. We want to find λ such that $\hat{u}$ is feasible for P_ε^- while its corresponding objective value is not so different from $c^T x_\varepsilon^+$.

Since for all $\|u\| \leq r_0$:

$$AP^\top(u_0 + u) = AP^\top u_0 + AP^\top u \leq b + \varepsilon,$$

then

$$AP^\top u_0 \leq b + \varepsilon - r_0 \begin{pmatrix} \|A_1 P^\top\| \\ \vdots \\ \|A_m P^\top\| \end{pmatrix}.$$

Therefore, we have, with probability $1 - \delta$,

$$AP^\top u_0 \leq b + \varepsilon - r_0(1-\varepsilon) \begin{pmatrix} \|A_1\| \\ \vdots \\ \|A_m\| \end{pmatrix} v = b + \varepsilon - r_0(1-\varepsilon).$$

Hence

$$AP^\top \hat{u} = (1-\lambda)AP^\top u_\varepsilon^+ + \lambda AP^\top u_0 \leq b + \varepsilon - \lambda r_0(1-\varepsilon) \leq b + \varepsilon - \frac{1}{2}\lambda r_0,$$

as we can assume w.l.o.g. that $\varepsilon \leq \frac{1}{2}$. Hence, $AP^\top \hat{u} \leq b$ if we choose $\varepsilon \leq \lambda \frac{r_0}{2}$. Moreover let $u \in B(u_0, r_0)$ such that u is collinear to u_0 and $\|u\| = r_0$, we have

$$\|u_0\| \leq \|u_0 + u\| - r_0 \leq 1 + \varepsilon - r_0,$$

so we have

$$\|\hat{u}\| \leq (1-\lambda)\|u_\varepsilon^+\| + \lambda\|u_0\| \leq (1-\lambda)(1+\varepsilon) + \lambda(1+\varepsilon-r_0) = 1+\varepsilon-\lambda r_0,$$

which is less than or equal to $1-\varepsilon$ for

$$\varepsilon \leq \frac{1}{2}\lambda r_0.$$

We can choose $\lambda = 2\frac{\varepsilon}{r_0}$. With this choice, $\hat{u}$ is a feasible point for the problem P_ε^-. Therefore, we have

$$c^\top P^\top u_\varepsilon^- \leq c^\top P^\top \hat{u} = c^\top P^\top u_\varepsilon^+ + \lambda c^\top P^\top (u_0 - u_\varepsilon^+) \leq c^\top P^\top u_\varepsilon^+ + \frac{4(1+\varepsilon)\varepsilon}{r_0}\|Pc\|.$$

By the above Lemma, we know that $r_0 \geq (1-\varepsilon)r^*$, therefore, we have:

$$c^\top P^\top u_\varepsilon^- \leq c^\top P^\top \hat{u} \leq c^\top P^\top u_\varepsilon^+ + \frac{4(1+\varepsilon)\varepsilon}{r_0}\|Pc\| \leq c^\top P^\top u_\varepsilon^+ + \frac{4(1+\varepsilon)^2\varepsilon}{(1-\varepsilon)r^*}\|c\|,$$

with probability at least $1-\delta$.

Theorem 6.2.6. *With probability at least $1-2\delta$, we have*

$$c^\top x_\varepsilon^+ \leq c^\top x_\varepsilon^- \leq c^\top x_\varepsilon^+ + \frac{18\varepsilon}{r^*}\|c\|.$$

Proof. It follows directly from the above discussions, with the notice that, when $0 \leq \varepsilon \leq \frac{1}{2}$, we can simplify:

$$\frac{2(1+\varepsilon)^2}{(1-\varepsilon)} \leq \frac{2(1+\frac{1}{2})^2}{(1-\frac{1}{2})} = 9.$$

□

6.2.3 Trust-region subproblems with quadratic models

In this subsection, we consider the case when the surrogate models using in TR methods are quadratic and defined as follows:

$$\min\{x^\top Q x \mid Ax \leq b,\ \|x\|_2 \leq 1, x \in \mathbb{R}^n\}. \tag{6.5}$$

Similar to the previous section, we also study the relations between this and two other problems:

$$\min\{u^\top PQP^\top u \mid AP^\top u \leq b,\ \|u\|_2 \leq 1-\varepsilon, u \in \mathbb{R}^d\} \tag{Q_ε^-}$$

and

$$\min\{u^\top PQP^\top u \mid AP^\top u \leq b+\varepsilon,\ \|u\|_2 \leq 1+\varepsilon, u \in \mathbb{R}^d\}. \tag{Q_ε^+}$$

We will just state the similar feasibility result in this case:

Theorem 6.2.7. *Let $P \in \mathbb{R}^{d\times n}$ be a random matrix in which each entry is an i.i.d $\mathcal{N}(0, \frac{1}{\sqrt{n}})$ random variable. Let $\delta \in (0,1)$. Assume further that*

$$n \geq \frac{(d+1)\log(\frac{2d}{\delta})}{c\varepsilon^2},$$

for some universal constant $c > \frac{1}{4}$. Then with probability at least $1-\delta$, for any feasible solution u of the projected problem (Q_ε^-), $P^\top u$ is also feasible for the original problem (6.5).

Let u_ε^- and u_ε^+ be optimal solutions for these two problems, respectively. Denote by $x_\varepsilon^- = P^\top u_\varepsilon^-$ and $x_\varepsilon^+ = P^\top u_\varepsilon^+$. Let x^* be an optimal solution for the original problem (6.1). We will try to bound $x^{*\top}Qx^*$ between $x_\varepsilon^{-\top}Qx_\varepsilon^-$ and $x_\varepsilon^{+\top}Qx_\varepsilon^+$, the two values that are expected to be approximately close to each other.

Theorem 6.2.8. *Let $P \in \mathbb{R}^{d\times n}$ be a random matrix in which each entry is an i.i.d $\mathcal{N}(0, \frac{1}{\sqrt{n}})$ random variable. Let $\delta \in (0,1)$. Assume further that*

$$n \geq \frac{(d+1)\log(\frac{2d}{\delta})}{C\varepsilon^2},$$

for some universal constant $C > \frac{1}{4}$ and

$$d \geq \mathcal{C}\frac{\log(m/\delta)}{\varepsilon^2},$$

for some universal constant $\mathcal{C} > 1$.
Let x^* be an optimal solution for the original problems (6.5). Then
(i) With probability at least $1-\delta$, the solution x_ε^- is feasible for the original problem (6.5).
(ii) With probability at least $1-\delta$, we have:

$$x_\varepsilon^{-\top}Qx_\varepsilon^- \geq x^{*\top}Qx^{*\top} \geq x_\varepsilon^{+\top}Qx_\varepsilon^+ - 3\varepsilon\|Q\|_*.$$

Proof. (i) From the previous theorem, with probability at least $1-\delta$, for any feasible point u of the projected problem (P_ε^-), $P^\top u$ is also feasible for the original problem (6.5). Therefore, it must hold also for x_ε^-.

(ii) From part (i), with probability at least $1-\delta$, x_ε^- is feasible for the original problem (6.5). Therefore, we have

$$x_\varepsilon^{-\top}Qx_\varepsilon^- \geq x^{*\top}Qx^*,$$

with probability at least $1-\delta$.

Moreover, due to Lemma 6.2.1, with probability at least $1-8ke^{-\mathcal{C}\varepsilon^2 d}$, where k is the rank of Q, we have

$$x^{*\top}Qx^* \geq x^{*\top}P^\top PQP^\top Px^* - 3\varepsilon\|x^*\|^2.\|Q\|_* \geq x^{*\top}P^\top PQP^\top Px^* - 3\varepsilon\|Q\|_*,$$

since $\|x^*\| \leq 1$. On the other hand, let $\hat{u} := Px^*$, due to Lemma 6.2.1, we have

$$AP^\top \hat{u} = AP^\top Px^* \leq Ax^* + \varepsilon \|x^*\| \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix} \leq Ax^* + \varepsilon \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix} \leq b + \varepsilon,$$

with probability at least $1 - 4me^{-C\varepsilon^2 d}$, and

$$\|\hat{u}\| = \|Px^*\| \leq (1+\varepsilon)\|x^*\| \leq (1+\varepsilon),$$

with probability at leats $1 - 2e^{-C\varepsilon^2 d}$ (this is the norm preservation property of random projections). Therefore, $\hat{u}$ is a feasible solution for the problem (P_ε^+) with probability at least $1 - (4m+2)e^{-C\varepsilon^2 d}$. Due to the optimality of u_ε^+ for the problem (P_ε^+), it follows that

$$\begin{aligned} x^{*\top} Q x^* &\geq x^{*\top} P^\top PQP^\top Px^* - 3\varepsilon \|Q\|_* \\ &= \hat{u}^\top PQP^\top \hat{u} - 3\varepsilon \|Q\|_* \\ &\geq u_\varepsilon^{+\top} PQP^\top u_\varepsilon^+ - 3\varepsilon \|Q\|_* \\ &= x_\varepsilon^{+\top} Q x_\varepsilon^+ - 3\varepsilon \|Q\|_*, \end{aligned}$$

with probability at least $1-(4m+6)e^{-C\varepsilon^2 d}$, which is at least $1-\delta$ for some universal constant C. □

The above result implies that the value of $x^{*\top} Q x^*$ lies between $x_\varepsilon^{-\top} Q x_\varepsilon^-$ and $x_\varepsilon^{+\top} Q x_\varepsilon^+$. It remains to prove that these two values are not so far from each other. For that, we also use the definition of fullness measure. We have the following result:

Theorem 6.2.9. *Let $0 < \varepsilon < 0.1$. Then with probability at least $1 - 2\delta$, we have*

$$x_\varepsilon^{+\top} Q x_\varepsilon^+ \leq x_\varepsilon^{-\top} Q x_\varepsilon^- < x_\varepsilon^{+\top} Q x_\varepsilon^+ + \frac{36\varepsilon}{\text{FULL}(S^*)} \|c\|.$$

Proof. Let $B(u_0, r_0)$ be a closed ball with maximum radius that is contained in S_ε^+.

We define $\hat{u} := (1-\lambda)u_\varepsilon^+ + \lambda u_0$ for some $\lambda \in (0,1)$ specified later. We want to find "small" λ such that $\hat{u}$ is feasible for Q_ε^- while its corresponding objective value is still close to $x_\varepsilon^{+\top} Q x_\varepsilon^+$.

Similar to the above proof, when we choose $\lambda := 2\frac{\varepsilon}{r_0}$, then $\hat{u}$ is feasible for the problem Q_ε^- with probability at least $1 - \delta$.

Therefore, $u_\varepsilon^{-\top} PQP^\top u_\varepsilon^-$ is smaller than or equal to

$$\begin{aligned} &\hat{u}^\top PQP^\top \hat{u} \\ &= \big(u_\varepsilon^+ + \lambda(u_0 - u_\varepsilon^+)\big)^\top PQP^\top \big(u_\varepsilon^+ + \lambda(u_0 - u_\varepsilon^+)\big) \\ &= u_\varepsilon^{+\top} PQP^\top u_\varepsilon^+ + \lambda u_\varepsilon^{+\top} PQP^\top \big(u_0 - u_\varepsilon^+\big) + \lambda(u_0 - u_\varepsilon^+)^\top PQP^\top u_\varepsilon^+ + \lambda^2 (u_0 - u_\varepsilon^+)^\top PQP^\top (u_0 - u_\varepsilon^+). \end{aligned}$$

However, from the Lemma 6.2.1 and the Cauchy-Schwartz inequality, we have

$$\begin{aligned} u_\varepsilon^{+\top} PQP^\top \left(u_0 - u_\varepsilon^+\right) &\leq \|P^\top u_\varepsilon^+\| \cdot \|Q\|_2 \cdot \|P^\top (u_0 - u_\varepsilon^+)\| \\ &\leq (1+\varepsilon)^2 \|u_\varepsilon^+\| \cdot \|Q\|_2 \cdot \|(u_0 - u_\varepsilon^+)\| \\ &\leq 2(1+\varepsilon)^4 \, \|Q\|_2 \\ &\quad \text{(Since } \|u_\varepsilon^+\| \text{ and } \|u_\varepsilon^-\| \leq 1+\varepsilon) \end{aligned}$$

Similar for other terms, then we have

$$\hat{u}^\top PQP^\top \hat{u} \leq u_\varepsilon^{+\top} PQP^\top u_\varepsilon^+ + (4\lambda + 4\lambda^2)(1+\varepsilon)^4 \, \|Q\|_2.$$

Since $\varepsilon < 0.1$, we have $(1+\varepsilon)^4 < 2$ and we can assume that $\lambda < 1$. Then we have

$$\begin{aligned} \hat{u}^\top PQP^\top \hat{u} &< u_\varepsilon^{+\top} PQP^\top u_\varepsilon^+ + 16\lambda \|Q\|_2 \\ &= u_\varepsilon^{+\top} PQP^\top u_\varepsilon^+ + \frac{32\varepsilon}{r_0} \qquad (\text{since } \|Q\|_2 = 1) \\ &\leq u_\varepsilon^{+\top} PQP^\top u_\varepsilon^+ + \frac{32\varepsilon}{(1-\varepsilon)\text{FULL}(S^*)} \qquad (\text{due to Lemma 6.2.5}) \\ &< u_\varepsilon^{+\top} PQP^\top u_\varepsilon^+ + \frac{36\varepsilon}{\text{FULL}(S^*)} \qquad (\text{since } \varepsilon \leq 0.1), \end{aligned}$$

with probability at least $1 - 2\delta$. The proof is done.

□

Chapter 7

Concluding remarks

7.1 Summary

This thesis focuses on some applications of random projections in optimization. It shows that random projections is a potential dimension reduction tool for many important optimization problems. In order to effectively apply this technique, the desired problem should be of really high dimension. Therefore, it is suitable for problems arising from "big-data" scenarios, such as those in machine learning and image processing. The thesis studies several other popular problems.

1. **Linear and Integer Programming.** We first use random projections to solve LPs and IPs in their standard form. Using bisection arguments, we can transform them into feasibility problems and apply random projections to their set of linear equality constraints. We show that the original and projected feasibility problems are strongly related (under some conditions) with high probability. Taking these results as a starting point, we then develop some algorithms to solve linear programming directly (without having to use binary search), thanks to the LP duality theory.

2. **Convex Programming with linear system of constraints.** We next consider the feasibility problem where a linear system is accompanied by other convex constraints. Similarly, the projected problem is formed by applying random projections to the set of linear equalities while keeping others unchanged. We consider two choices of random projections: one is sampled from a sub-gaussian distributions, one is from a randomized orthogonal system. The relations between the two problems are established through the so-called "Gaussian width".

3. **Membership problem.** We extend our study to the general case of membership problem, in which we ask whether a given point belongs to a given set . In this case, we prove that if we reach the "true dimension" of the set, we can always separate the projected point from the projected set. In order to do that, we use the concept of "doubling dimension" to quantify the intrinsic dimension of a point set. We also generalize this result to the case when a threshold distance between a point and a set is required.

4. **Trust-region subproblem**. Lastly, we apply random projections to study trust-region subproblems. This is the only time we are able to reduce the number of variables but not the number of constraints. We prove that the linear and quadratic models can be well approximated by those in a lower projected dimension. We quantify the error between the two problems using the so-called "fullness measure" of a set. The results suggest that random projection can be used to study high dimensional derivative-free optimization problems, which is intractable at the current time.

7.2 Further research

- It should be noted that, if a problem consists of linear constraints, we can directly apply a random projection to them. Therefore, we are able to reduce the number of (linear) constraints. On the other hand, if those linear constraints are written in the inequality form, then we can reduce the number of variables (as in the case of the trust-region subproblems in Chapter 6). However, we can not reduce both of them. One of our future researches therefore will focus on reducing both these two quantities (at the same time): the number of variables and the number of constraints. Our intuition tells us that using random projections alone will not work, and we need to combine random projections with other dimension reduction techniques. At the moment, we are looking at the multiplicative weight update (MWU) algorithm, which seems to be a potential candidate.

- Another research direction would be to extend the results and techniques developed in this thesis to study other different problems. Currently, we are studying the use of random projection and measure concentration in semidefinite programming (SDP). We also planned to look at several problems arising from "machine learning" such as constrained linear regression and support vector machine. They seem to be very amenable to random projection techniques.

- One of the weaknesses of this thesis is the lack of numerical tests and "real-work appli-

cations". We will try to improve this facet by looking at more practical problems with similar form as those being studied here. At the moment, we are studying the "quantile regression", a problem that can be reformulated as a (dense) linear feasibility problem. The preliminary computational results are quite promising. It seems to be the right application to illustrate the success of random projections.

- In addition to reducing dimension, we also want to investigate some side effects of random projections. For example, they turn a very ill-scaled system (of linear equalities) into a good one. Therefore, we want to know how the condition number of the projected matrix TA changes, compared to the condition number of the original matrix A.

Bibliography

[1] Dimitris Achlioptas. Database-friendly random projections: Johnson-Lindenstrauss with binary coins. *Journal of Computer and System Sciences*, 66(4):671 – 687, 2003. Special Issue on {PODS} 2001.

[2] P. Agarwal, S. Har-Peled, and H. Yu. Embeddings on surfaces, curves, and moving points in Euclidean space. In *Proceedings of the 23rd Symposium on Computational Geometry*, pages 381–389. ACM, 2007.

[3] Nir Ailon and Bernard Chazelle. The fast Johnson-Lindenstrauss transform and approximate nearest neighbors. *SIAM Journal on Computing*, 39(1):302–322, 2009.

[4] Nir Ailon and Edo Liberty. Fast dimension reduction using Rademacher series on dual BCH codes. *Discrete & Computational Geometry*, 42(4):615–630, 2009.

[5] N. Alon. Problems and results in extremal combinatorics - I. *Discrete Mathematics*, 273:31–53, 2003.

[6] Ella Bingham and Heikki Mannila. Random projection in dimensionality reduction: Applications to image and text data. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '01, pages 245–250, New York, NY, USA, 2001. ACM.

[7] Emmanuel J. Candès. Mathematics of sparsity (and few other things). In *Proceedings of the International Congress of Mathematicians*. Seoul, South Korea, 2014.

[8] Sanjoy Dasgupta and Yoav Freund. Random projection trees and low dimensional manifolds. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 537–546, New York, NY, USA, 2008. ACM.

[9] Sanjoy Dasgupta and Anupam Gupta. An elementary proof of a theorem of Johnson and Lindenstrauss. *Random Struct. Algorithms*, 22(1):60–65, January 2003.

[10] L. Gottlieb and R. Krauthgamer. Proximity algorithms for nearly doubling spaces. *SIAM Journal on Discrete Mathematics*, 27(4):1759–1769, 2013.

[11] Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '03, pages 534–, Washington, DC, USA, 2003. IEEE Computer Society.

[12] P. Indyk and A. Naor. Nearest neighbor preserving embeddings. *ACM Transactions on Algorithms*, 3(3):Art. 31, 2007.

[13] Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: Towards removing the curse of dimensionality. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, pages 604–613, New York, NY, USA, 1998. ACM.

[14] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. *Contemporary Mathematics*, 26:189–206, 1984.

[15] R. Krauthgamer and J.R. Lee. Navigating nets: Simple algorithms for proximity search. In *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 791–801, 2004.

[16] Kasper Green Larsen and Jelani Nelson. The Johnson-Lindenstrauss lemma is optimal for linear dimensionality reduction. *CoRR*, abs/1411.2404, 2014.

[17] A. Magen. Dimensionality reductions in ℓ_2 that preserve volumes and distance to affine spaces. *Discrete and Computational Geometry*, 30(1):139–153, 2007.

[18] Jiří Matoušek. On variants of the Johnson-Lindenstrauss lemma. *Random Struct. Algorithms*, 33(2):142–156, September 2008.

[19] Jiří Matoušek. Lecture notes on metric embeddings. *Manuscript*, 2013.

[20] A. Mood, F. Graybill, and D. Boes. *Introduction to the Theory of Statistics*. McGraw-Hill, 1974.

[21] Mert Pilanci and Martin J. Wainwright. Randomized sketches of convex programs with sharp guarantees. *CoRR*, abs/1404.7203, 2014.

[22] Mert Pilanci and Martin J. Wainwright. Randomized sketches of convex programs with sharp guarantees. *IEEE Trans. Information Theory*, 2015.

[23] J.-L. Verger-Gaugry. Covering a ball with smaller equal balls in $\mathbb{R}^n$. *Discrete Computational Geometry*, 33:143–155, 2005.

[24] Ky Vu. Randomized sketches for convex optimization with linear constraints. *Preprint*, 2015.

[25] Ky Vu, Pierre-Louis Poirion, and Leo Liberti. Gaussian random projections for Euclidean membership problems. *arXiv preprint arXiv:1509.00630*, 2015.

[26] Ky Vu, Pierre-Louis Poirion, and Leo Liberti. Using the Johnson-Lindenstrauss lemma in linear and integer programming. *arXiv preprint arXiv:1507.00990*, 2015.

[27] Ky Vu, Pierre-Louis Poirion, and Leo Liberti. Random projections for trust-region subproblems with applications to derivative-free optimization. *Preprint*, 2016.

[28] L. Zhang, M. Mahdavi, R. Jin, and T. Yang. Recovering optimal solution by dual random projection. In *Conference on Learning Theory (COLT) JMLR W & CP*, volume 30, pages 135–157, 2013.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Projection aléatoire pour l'optimisation de grande dimension

Mots clés : Réduction de dimension , algorithme aléatoire , Johnson - Lindenstrauss lemme

Résumé : Les projection aléatoires sont une technique très utile pour réduire la dimension des données et a été largement utilisé dans l'algèbre linéaire numérique, le traitement d'image, de l'informatique, l'apprentissage machine et ainsi de suite. Une projection aléatoire est souvent définie comme une matrice aléatoire construite d'une certaine manière de telle sorte qu'elle conserve de nombreuses propriétés importantes, y compris les distances, les produits internes, les volumes ... de l'ensemble de données. L'un des exemples les plus célèbres est le lemme de Johnson-Lindenstrauss, qui affirme qu'un ensemble de m points peut être projeté, par une projection aléatoire, dans un espace euclidien de dimension O (log m) tout en assurant que les distances entre les points reste environ inchangé..

Dans cette thèse, nous appliquons des projections aléatoires pour étudier un certain nombre de problèmes d'optimisation importants comme la programmation linéaire et la programmation en nombres entiers, les problèmes d'appartenance convexes et l'optimisation des produits dérivés libre. Nous nous sommes particulièrement intéressés aux cas où les dimensions du problème sont si élevés que les méthodes traditionnelles ne peuvent pas être appliquées. Dans ces conditions, au lieu de traiter directement avec les problèmes d'origine, nous appliquons des projections aléatoires pour les transformer en des problèmes de dimensions beaucoup plus faibles. Nous montrons que, tout en obtenant des problèmes beaucoup plus facile à résoudre, ces nouveaux problèmes sont de très bonnes approximations des originaux. Cela permet de suggérer que les projections aléatoires sont un outil très prometteur de réduction de dimension pour beaucoup d'autres problèmes..

Title : Random projection for high dimensional optimization.

Keywords : Dimension reduction, randomized algorithm, Johnson-Lindenstrauss lemma

Résumé : Random projection is a very useful technique for reducing data dimension and has been widely used in numerical linear algebra, image processing, computer science, machine learning and so on. A random projection is often defined as a random matrix constructed in certain ways such that it preserves many important features, including distances, inner products, volumes… of the data set. One of the most famous examples is the Johnson-Lindenstrauss lemma, which asserts that a set of m points can be projected by a random projection, to an Euclidean space of dimension O(log m) whilst still ensures that the inner distances between them approximately unchanged.

In this PhD thesis, we apply random projections to study a number of important optimization problems such as linear and integer programming, convex membership problems and derivative-free optimization. We are especially interested in the cases when the problem dimensions are so high that traditional methods can not be applied. In those circumstances, instead of dealing directly with the original problems, we apply random projections to transform them into much lower dimensional problems. We prove that, while getting much easier to solve, these new problems are very good approximation of the original ones. It is suggested that random projection is a very promising dimension reduction tool for many other problems as well.catervis mixtae praedonum.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France