# Review 41: Improving GAN Training with Probability Ratio Clipping and Sample Reweighting

**Paper: https://arxiv.org/abs/2006.06900v4**

פינת הסוקר:

המלצת קריאה ממייק: מומלץ אך לא חובה לאלו שרוצים להתעמק בשיטות אימון של GANs.

בהירות כתיבה: בינונית פלוס.

רמת היכרות עם כלים מתמטיים וטכניקות של ML/DL הנדרשים להבנת מאמר: הבנה טובה ווסרשטיין גאן וכל מה שקשור אליו, הכרה בסיסית בשיטות מעולם הסטטיסטיקה כמו importance sampling, רקע בסיסי בלמידה באמצעות חיזוקים (Reinforcement learning) .

יישומים פרקטיים אפשריים: אימון גאן משופר במגוון תרחישים

פרטי מאמר:

פורסם בתאריך: 30.10.2020, בארקיב.

הוצג בכנס: NeurIPS 2020.

תחומי מאמר:

גאנים.

שיטות אימון של גאנים.

כלים מתמטיים, מושגים וסימונים:

וסרשטיין WGAN) GAN).

מרחק וסרשטיין (WD).

פונקצית ליפשיץ.

שיטות וריאציוניות לבעיות אופטימיזציה בתחום הרשתות הגנרטיביות כמו GAN.

גישות מתורת למידת החיזוק (RL): אופטימיזציה של פוליסי (Policy Optimization - PO) דרך פתרון של בעיית אופטימיזציה עם פונקצית מטרה חלופית - surrogate.

שיטות דגימה: IM) Importance Sampling).

מרחקים בין מידות הסתברות: מרחק KL ומרחק KL הפוך.

אלגוריתמים של EM) Expectation-Maximization).

תמצית מאמר:

אתם בטח יודעים שלמרות מאמצי מחקר אינטנסיביים בשנים האחרונות, האימון של GAN-ים עלול להוות משימה לא טריוויאלית עקב קושי במציאת איזון בין הגנרטור G לדיסקרימינטור D. המאמר הנסקר מציין שבעיות אלו בולטות במיוחד בתחום גנרוט טקסט עקב האופי הדיסקרטי של משימה זו (נציין שכרגע שיטות SOTA למשימות גנרוט של טקסט אינם מבוססות על GAN-ים). כדי להתגבר על סוגיות אלו, מאמר הנסקר מציע שיטה לשיפור תהליך האימון של GAN שמבוססת על שני רעיונות עיקריים:

מניעה עדכונים גדולים מדי של הגנרטור G שעלולים לפגוע ביציבות של תהליך האימון ולהוביל לאובדן של איזון בין G לדיסקרימינטור D. איזון זה הינו חיוני להתכנסות של תהליך האימון של GAN ולפתרון איכותי עבור בעיית אופטימיזציה מינימקס ש-GAN מנסה לפתור. נזכיר שתהליך אימון של GAN הינו משחק סכום אפס כאשר G מאומן לגרום ל-D לזהות את הדאטה הסינטטי ש-G מייצר כדאטה אמיתי (מסט האימון) ובתורו D מאומן להבחין בין דגימות ש-G מייצר לאמיתיות.

משקול של דגימות המגונרטות עי" G בתהליך האימון של D. כאמור D מאומן להבחין בין דגימות אמיתיות (מאומן לתת ציון גבוה) מסט האימון לבין דגימות המגונרטות עי" G (מאומן לתת ציון נמוך). בתהליך עדכון של D הדגימות של G באיכות טובה שמצליחות "לעבוד יותר טוב על D" (בעלי ציון גבוה) מקבלות משקל גבוה ואילו דגימות של G ה "פחות אמיתיות" מבחינת D (בעלי ציון נמוך) מקבלות משקל נמוך נמוך יותר. זה הופך את האימון של D ליעיל יותר כי (לטענת המאמר) הוא לא מתבזבז על עדכונים על דגימות קלות מדי (האינטואיציה כאן אומרת שאם D משקיע מאמץ רב יותר בלהתאמן על דגימות איכותיות יותר, הוא יהיה מספיק חזק בשביל להפגין ביצועים טובים גם על דגימות קלות יותר ב"צורה אוטומטית").

הערה: גישה זה מזכירה לי שיטות ממשפחת GBM) gradient boosting machines) מממשקלות דוגמאות בהתאם ל"רמת הקושי" שלהם מבחינת המודל (בגדול עד כמה השערוך של המודל מדויק).

הסבר של רעיונות בסיסיים:

וסרשטיין GAN: נקודת ההתחלה של המאמר זה WGAN, המודיפיקציה של ה-GAN המקורי, המשתמשת במרחק וסרשטיין (WD) כבסיס ל-D. כלומר G מאומן לגנרט דגימות בעלות מרחק וסרשטיין נמוך מהדוגמאות מסט האימון. מרחק וסרשטיין הינו מקרה פרטי של טרנספורט אופטימלי וכבר הסברתי על באחד הפוסטים שלי (Learning to summarize from human feedback).

היתרון הבולט של WGAN על GAN רגיל טמון ביכולת של D "להעביר גרדיאנטים" יותר יציבים ל-G גם במקרים כאשר D מצליח בקלות להבדיל בין הדגימות האמיתיות לדגימות המגונרטות. זה קורה בגלל שלהבדיל ממרחק JS) Jensen-Shannon) שאותו מנסה למזער ה-GAN הרגיל, WD הינו בעל אופי רציף יותר ולא מגיע לרוויה (כמו מרחק JS) גם כאשר התפלגות הדגימות של G רחוקה מאוד מההתפלגות של הדאטה סט (המשוערכת ע"י D).

חישוב של מרחק וסרשטיין לפי הגדרתו הינו משימה מאוד קשה ובדרך כלל פותרים את בעיית האופטימיזציה הדואלית שלה (שוויון רובינשטיין-קנטורוביץ'). הבעיה הדואליות הינה המקסום של הפרש התוחלות בין התפלגויות של דאטה האמיתי לבין הדגימות המגונרטות מעל מרחב של פונקציות k-ליפשיץ רציפה, מוכפלת ב 1/k. פונקציה זו ממודלת עי" רשת נוירונים כאשר נעשים טריקים שונים, כמו קיצוץ משקלים או אילוצים על הנגזרת של הפונקציה כדי שהפונקציה הממודלת תהיה k-ליפשיץ רציפה). אז בעיית אופטימיזציה ש- WGAN מנסה לפתור, הינה מקסום של הפרש התוחלות זה מעל מרחב כל פונקציות k-ליפשיץ רציפות f, מבחינת D. הגנרטור G מצידו מנסה למזער אותו הפרש התוחלות המתואר לעיל (בעיית מינימקס). אם נתבונן בפונקציית מטרה של WGAN ניתן לראות כי G מנסה למקסם את התוחלת של פונקצית ליפשיץ f (על מרחב הדגימות שלו). ניתן למצוא דמיון בין בעיית אופטימיזציה זו לבין אופטימיזציה של פוליסי בעולם של RL, כאשר פונקציה k-ליפשיץ רציפה f משחקת תפקיד של גמול (reward) והתפלגות דגימות של G ניתן לראות כפוליסי. דמיון זה, שזוהה בכמה מאמרים של השנים האחרונות, ינוצל בבניה של פונקצית מטרה חדשה ל WGAN שהוצעה במאמר.

אחרי שהבנו מה זה WGAN ואת הקשר שלו לבעיות RL, בואו נתקדם בשינוי של פונקציית מטרה של WGAN המוצע עי" המאמר. פתרונה יוביל למניעה של עדכונים גדולים של G ומשקול דגימות, המבוסס על ה"איכות" שלהן בעדכונים של D. לאור הקשר עם בעיות של אופטימיזציה של פוליסת ב-RL, השיטה שהמאמר מציע דומה לשיטות של אופטימיזציה של פוליסי כמו PPO ו- TRPO. שיטות אלה מחליפות את פונקצית המטרה הרגילה בפונקציה חלופית שמנסה לשפר את פונקציית הפוליסי F\_p. זה נעשה עי״ מקסום התוחלת של פונקצית היתרון המוכפלת ביחס של F\_p החדשה ל- F\_p הישנה תחת אילוץ שמרחק KL בין F\_p החדשה לישנה חסום עי״ קבוע קטן (אילוץ זה מופיע לפעמים האיבר רגולריזציה בפונקצית המטרה). בדרך זו F\_p החדשה לומדת לתת הסתברויות גבוהות למצבים שבהם פונקצית היתרון מקבלת ערכים גבוהים כלומר הגמול אחרי עדכון של P\_i הינו מקסימלי).

פונקציית המטרה של המאמר: המאמר מציע להחליף את פונקציית המטרה הסטנדרטית של WGAN בפונקציה F\_imp המכילה הפרש של שני האיברים הבאים:

איבר 1: התוחלת של פונקציה k-ליפשיץ רציפה f מעל מידת הסתברות עזר q (שתלויה בהתפלגות הדגימות המגונרטות P\_g וגם בפונקצית f הממודלת עי" D בצורה מפורשת ולא פרמטרית (!!!)).

איבר 2: מרחק KL בין q לבין P\_g.

המאמר מציע לאמן את WGAN עי" מקסום של F\_imp, כאשר הפרמטרים הם משקלי הרשתות של G ו-D. אם נזכר בעובדה שמרחק KL הינו תמיד אי שלילי, ניתן להבין שהמקסום של F\_imp שקול למקסום של האיבר הראשון המינימיזציה של האיבר השני. אז ניתן לפרש את בעיית מקסום F\_imp באופן הבא:

מקסום של תוחלת הציון הניתן עי" D להתפלגות q (האיבר הראשון) כאשר אנו מנסים לשמור את התפלגות הדגימות של G קרובה ל q.

אימון של G: מקסום של F\_imp מבחינת הפרמטרים של G, הינו מקרה קלאסי של בעיית אינפרנס ורציאונית שמזכירה את בעיית אופטימיזציה שאנו פותרים למשל בVAE- Variational AutoEncoder. הדרך הטבעית לפתור אותה הינה להשתמש באלגוריתם EM קלאסי. בשלב E של EM, אנו מוצאים את ההתפלגות g שהיא בצורה של מכפלה של אקספוננט של P\_g ושל f (מנורמלת). שימו לב שמה שיש מכפלה זו מהווה משקול של P\_g, כאשר הדגימות עם ציון של D יותר גבוה מקבלות הסתברות גבוהה יותר, שזה מה שרצינו מההתחלה.

השלב M של האלגוריתם הינו אופטימיזציה של F\_imp על הפרמטרים של G כאשר התפלגות q נתונה (חושבה בשלב E). זה למעשה מינימיזציה של האיבר השני, מרחק KL. וכאן יש לנו בעייה כי q זה בעצם פונקציה של P\_g הניתנת בצורה לא מפורשת ובשביל לשערך את מרחק KL נצטרך לדגום מ-q שזה מאוד לא טריוויאלי. למזלנו ניתן להשתמש ב-KL הפוך ולהפוך את האיבר זה לסכום של מינוס התוחלת של f מעל P\_g ומרחק KL בין P\_g עבור האיטרציה הקודמת לבין P\_g שאנו מנסים לאפטם (נוסחה 4 במאמר). בעצם אנו מנסים למקסם את התוחלת של f מעל P\_g אך לא רוצים להתרחק מדי מההתפלגות P\_g מהאיטרציה הקודמת. אם אתם זוכרים את ההסבר שלי על PPO ועל TRPO, מיד תזהו את הדמיון. אז בדומה לשיטות אלו, המאמר מציע להחליף את פונקציית המטרה כאן בפונקציית מטרה חלופית המכילה המכפלה של פונקציה f ביחס בין P\_g הישן לחדש r\_g(!!). בנוסף הם מאלצים את r\_g להיות קטן באופן מאולץ (מקצצים). אבל כאן יש לנו עוד בעיה. איך נחשב את היחס הזה על דגימה של G אם P\_g נתון בצורה לא מפורשת. כאן הם עושים טריק נחמד. בנוסף ל D של WGAN, הם מאמנים דיסקרימינטור בינארי D\_bin בשביל להבדיל בין הדגימות של G לדגימות האמיתיות. ניתן להוכיח (עשו זאת במאמר המקורי של GAN למשל) שעבור D\_bin אופטימלי ניתן לחשב את ערך של P\_g עבור הדגימה של הערך של D\_bin הדגימה זו. בדרך זו ניתן לשערך את r\_g עבור דגימה נתונה.

אימון של D: כאן אנו צריכים לאפטם רק את האיבר הראשון (התוחלת של f מעל התפלגות q נתון כאשר מאפטמים את הפרמטרים של f). כאן משתמשים כמובן ב Gradient Descent אבל נשאלת השאלה איך נחשב את הגרדיאנט עבור הפרמטרים של f אם אנחנו לא יודעים לדגום מ-q. בשביל להתגבר על הקושי הזה הם משתמשים בטכניקה קלאסית בסטטיסטיקה הנקראת IM תוך ניצול של הצורה של q (מכפלה של אקספוננט של P\_g ושל f). בתור התפלגות proposal שדוגמים ממנו במקום q, הם לקחו את P\_g שקל לדגום ממנה. נציין שהתוחלת של הגרדיאנט מעל q של f יוצאת שווה לתוחלת מעל P\_g של המכפלה של f באקספוננט של f. כך אנו משיגים את המשקול הגבוה לדגימות בעלות ציון גבוה מ D משפיעות יותר חזק על העדכון של D כאשר השפעה של דגימות עם ציון נמוך על עדכון של D קטנה (!!).

הישגי מאמר:

דומיין של תמונות: המאמר מראה שהשיטה שלהם משפרת את איכות התמונות מבחינת Inception Score ו- Frechet Distance מול כמה GAN-ים וביניהם אלו המבוססים על הלוס של WGAN עם טכניקות ייצוב אימון שונות וגם על כמה GAN-ים עם פונקציות לוס אחרת (לא בסגנון וסרשטיין). הם גם מראים שהם אכן מצליחים לייצב את האימון ועבור WGAN קלאסי (השונות של גרדיאנטים נמוכה יותר וההתכנסות יותר מהירה). הניסויים נעשו בעיקר על CIFAR10.

דומיין טקסטואלי: הם הצליחו לשפר את איכות הטקסט המגונרט - ההשוואה נעשתה עי״ BLEU. מעניין שהם גם הצליחו לשפר את איכות ביצוע המשימה של ״העברת סגנון״ (Style Transfer) כאשר המטרה כאן לשנות את סגנון המשפט (למשל סנטימנט) תוך כדי שימור התוכן.

נ.ב.

אחד המאמרים היפים מבחינת האלגנטיות המתמטית המתבטא השילוב טכניקות מתחומים שונים (לא ציינתי בסקירה שהם מוכיחים שהגישה שלהם מקדמת את ההתפלגות של G לכיוון של התפלגות הדאה האמיתית). לגבי הישימות של גישה זו חייבים לבחון אותה על דאטה סטים יותר מגוונים ועל משימות מורכבות יותר.