לילה טוב חברים, היום אנחנו שוב בפינתנו deepnightLearners עם סקירה של מאמר בתחום הלמידה העמוקה.
היום בחרתי לסקירה את המאמר שנקרא:
Improving GAN Training with Probability Ratio Clipping and Sample Reweighting
שיצא לפני כ 3 שבועות

תחומי מאמר: רשתות גנרטיביות, שיטות אימון של (Generative Adversarial Nets (GAN

כלים מתמטיים, מושגים וסימונים:

- GAN
- וסרשטיין GAN (WGAN), מרחק וסרשטיין(WD), פונקצית ליפשיץ
- שיטות וריאציוניות לבעיות אופטימיזציה בתחום הרשתות הגנרטיביות כמו GAN
- <u>גישות מתורת למידת החיזוק (:RL)</u> אופטימיזציה של פוליסי (Policy Optimization - PO) דרך פתרון של בעיית אופטימיזציה עם פונקצית מטרה חלופית - surrogate
- <u>שיטות דגימה</u>(Importance Sampling (IM :
- <u>מרחקים בין מידות הסתברות</u>: מרחק KL ומרחק KL הפוך
- אלגוריתמים של Expectation-Maximization (EM)
- גרדיינט דסצנט (GD)

תמצית מאמר: אתם בטח יודעים שלמרות מאמצי מחקר אינטנסיביים בשנים האחרונות, האימון של GANים עלול להוות משימה לא טריוויאלית עקב הקושי במציאת איזון בין הגנרטור G לדיסקרימינטור D. המאמר מציין שבעיות אלו בולטות במיוחד בתחום גנרוט טקסט עקב האופי הדיסקרטי שלו(נציין שכרגע שיטות SOTA למשימות גנרוט של טקסט <u>אינם</u> מבוססות עלGAN ים). כדי להתגבר על סוגיות אלו, מאמר זה מציע שיטה לשיפור תהליך האימון של GAN הבנויה על שני רעיונות עיקריים:

- <u>מניעה עדכונים גדולים מדי של הגנרטור G</u> שעלולים לפגוע ביציבות של תהליך האימון ועלול להוביל 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 בשביל לגרום ל D לזהות את הדגימות שלו כדגימות מהדאטה סט ובתורו D מאומן להבחין בין דגימות של G לאמיתיות.
- <u>משקול של דגימות המגונרטות עי" G בתהליך האימון של D</u>. כמו שאתם זוכרים D מאומן להבחין בין דגימות אמיתיות (מאומן לתת ציון גבוה) מהדאטה סט לבין דגימות המגונרטות עי" G (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").
<u>הערה</u>: גישה זה מזכירה לי שיטות ממשפחת gradient boosting machines (GBM) מממשקלות דוגמאות בהתאם ל"רמת הקושי" שלהם מבחינת המודל (בגדול עד כמה השערוך של המודל מדויק).

הסבר של רעיונות בסיסיים:
<u>וסרשטיין GAN</u>: נקודת ההתחלה של המאמר זה WGAN, המודיפיקציה של ה- GAN המקורי, המשתמשת במרחק וסרשטיין (WD) כבסיס ל D. כלומר G מאומן לגנרט דגימות עם מרחק וסרשטיין מינימלי מהדאטה סט. הינו מקרה פרטי של טרנספורט אופטימלי וכבר הסברתי על באחד הפוסטים שלי (https://www.facebook.com/groups/MDLI1/permalink/1724336801063694).

היתרון הבולט של WGAN על GAN רגיל הוא יכולת 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 Jensen-Shannon(JS) שאותו מנסה למזער הGAN הרגיל, WD הינו בעל אופי רציף 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 על בין התפלגויות של דאטה האמיתי לבין הדגימות המגונרטות מעל מרחב של פונקציות ליפשיץ עם מקדם 1(פונקציה זו ממודלת עי" רשת נוירונים כאשר נעשים טריקים שונים, כמו קיצוץ משקלים או אילוצים על הנגזרת של הפונקציה כדי שהפונקציה הממודלת תהיה ליפשיצית עם קבוע 1). אז בעיית אופטימיזציה ש- WGAN מנסה לפתור הינה מקסום של הפרש התוחלות זה על מרחב פונקציות ליפשיץ מבחינת D, כאשר G מצידה מנסה למזער אותו (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 f משחק 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 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: אז המאמר מציע להחליף את פונקציית המטרה הסטנדרטית של WGAN בפונקציה F_imp המכילה הפרש של שני האיברים הבאים:

- איבר 1: התוחלת של פונקציית ליפשיץ 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, קל 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 W_imp מבחינת הפרמטרים של G, הינו מקרה קלאסי של בעיית אינפרנס ורציאונית שמזכירה את בעיית אופטימיזציה שאנו פותרים למשל ב VAE - 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 W_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 GD 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 ממנו. 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 GANים וביניהם אלו המבוססים על הלוס של WGAN עם טכניקות ייצוב אימון שונות וגם על כמה GANים 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.

לינק למאמר: https://arxiv.org/abs/2006.06900

לינק לקוד: https://github.com/Holmeswww/PPOGAN

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.