Proof of Convergence of FedGAN

Firstly, the convergence of GAN is discussed [1] and then, the convergence of FL is discussed [2].

1) Convergence of GAN: The notations to be used in the discussion of convergence in GAN are defined as follows:

- Data and latent spaces: X, Z.
- Distributions: p_{data} (true data), p_z (latent prior), p_{θ_G} (generator-induced).
- Networks and parameters: $G_{\theta_G} : Z \to X$, $D_{\theta_D} : X \to \mathbb{R}$, with $\theta_G \in \mathbb{R}^{m_G}$, $\theta_D \in \mathbb{R}^{m_D}$.
- Loss link: $f : \mathbb{R} \to \mathbb{R}$ concave ($f(x) = -\log(1+e^{-x})$ for classical GAN).
- Equilibrium: $(\theta_D^\star, \theta_G^\star)$ with $\nabla_{\theta_D} V = \nabla_{\theta_G} V = 0$.
- Supports and neighborhoods: $\text{supp}(p)$, Euclidean ball $B_\varepsilon(\cdot)$.
- Curvature matrices:

$$K_{DD} = \mathbb{E}_{x\sim p_{\text{data}}}\big[\nabla_{\theta_D} D(x)\, \nabla_{\theta_D} D(x)^\top\big],$$

$$K_{DG} = \int_X \nabla_{\theta_D} D(x)\, \nabla_{\theta_G}^\top p_{\theta_G}(x)\, dx$$

- Spectral notation: $\lambda_{\max}(\cdot)$, $\lambda_{\min}^{(+)}(\cdot)$ (smallest nonzero eigenvalue).
- "Hurwitz" matrix: all eigenvalues have strictly negative real parts.
- State vector and field: $\theta \triangleq (\theta_D, \theta_G)$, $h(\theta) \triangleq (\dot{\theta}_D, \dot{\theta}_G)$.

The GAN objective as well as the general formulation of GAN optimization is presented by the following min-max problem [1]:

$$\begin{aligned} minmax V(D,G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[f(D_{\theta_D}(x))] \\ + \mathbb{E}_{z\sim p_z(z)}[f(-D_{\theta_D}(G_{\theta_G}(z)))] \end{aligned} \tag{1}$$

where

$$f(D_{\theta_D}(x)) = \log D_{\theta_D}(x) \tag{2}$$

$$f(D_{\theta_D}(G_{\theta_G}(z))) = \log(1 - D_{\theta_D}(G_{\theta_G}(z))) \tag{3}$$

By considering that the gradient steps in both θ_D and θ_D are simultaneous, the following equation is obtained:

$$\dot{\theta}_D = \nabla_{\theta_D} V(\theta_G, \theta_D), \qquad \dot{\theta}_G := \nabla_{\theta_G} V(\theta_G, \theta_D) \tag{4}$$

The authors in [1] demonstrated the local stability of the general GAN systems and stated Theorem 1, under the following assumptions [1]:

- **Assumption I (Good Equilibrium; realizable case)**

$$p_{\theta_G^\star} = p_{\text{data}}, \qquad D_{\theta_D^\star}(x) = 0,\ \forall x \in \text{supp}(p_{\text{data}}) \tag{5}$$

 with the intuition that the discriminator is "indifferent" at the true generator, and real and generated distributions locally coincide. A non-realizable variant replaces equality by local indistinguishability on the support.

- **Assumption II (Concavity at zero)**

$$f''(0) < 0, \qquad f'(0) \neq 0 \tag{6}$$

 This assumption ensures strict local concavity along the discriminator direction and non-degenerate coupling to the generator.

- **Assumption III (Curvature/identifiable)** The Hessians, at equilibrium, of two auxiliary convex quantities— (i) the discriminator deviation $\mathbb{E}[D^2_{\theta_D}(x)]$ (function of θ_D), and (ii) the squared magnitude of the discriminator update at $\theta_D^\star$ as a function of θ_G— are positive semi-definite and satisfy a "flat-only-if-constant" property along any direction of zero curvature. This aligns zero-curvature directions with a local subspace of equilibria.
- **Assumption IV (Support stability)**

$$\exists\, \varepsilon_G > 0 :\ \forall\, \theta_G \in B_{\varepsilon_G}(\theta_G^\star), \quad \text{supp}(p_{\theta_G}) = \text{supp}(p_{\text{data}}) \tag{7}$$

 This eliminates spurious support shifts that might break the equilibrium structure.

Under Assumptions I–IV, if $K_{DD} \succ 0$ and K_{DG} has full column rank, the Jacobian J of the system is Hurwitz and at equilibrium it is presented by the following equation:

$$J = \begin{bmatrix} 2f''(0)\, K_{DD} & f'(0)\, K_{DG} \\ -f'(0)\, K_{DG}^\top & 0 \end{bmatrix}$$

Based on the Assumptions I, II, III, and IV, Theorem 1 is stated as follows:

Theorem 1: With respect to an equilibrium point $(\theta_D^\star, \theta_G^\star)$, when Assumptions I, II, III, and IV hold for $(\theta_D^\star, \theta_G^\star)$ and other equilibria in a small neighbouhood, the dynamic system defined by the GAN objective stated in equation(1) and the updates in equation (4) is locally exponentially stable. The convergence rate is controlled only by the eigenvalues λ of the Jacobian J of the system at equilibrium with a strict negative real part upper bounded as:

- If $\text{Im}(\lambda) = 0$,

$$\text{Re}(\lambda) \leq \frac{A}{Y} \tag{8}$$

 where

$$A = 2f''(0)\, f'(0)^2\, \lambda_{\min}^{(+)}(K_{DD})\, \lambda_{\min}^{(+)}(K_{DG}^\top K_{DG})$$

 and

$$\begin{aligned} Y = 4f''(0)^2\, \lambda_{\min}^{(+)}(K_{DD})\, \lambda_{\max}(K_{DD}) \\ + f'(0)^2\, \lambda_{\min}^{(+)}(K_{DG}^\top K_{DG}) \end{aligned}$$

- If $\text{Im}(\lambda) \neq 0$,

$$\text{Re}(\lambda) \leq f''(0)\, \lambda_{\min}^{(+)}(K_{DD}) < 0. \tag{9}$$

The matrices of the form $\begin{bmatrix} -Q & P \\ -P^\top & 0 \end{bmatrix}$ are Hurwitz when $Q \succ 0$ and P has full column rank. Setting $Q = -2f''(0)K_{DD} \succ 0$ (since $f''(0) < 0$, $K_{DD} \succ 0$) and $P = f'(0)K_{DG}$ yields the claim. Exponential decay follows from a quadratic Lyapunov function $V(\delta) = \delta^\top P \delta$ with $J^\top P + PJ = -Q$. The discriminator block J_{DD} supplies dissipation (strict negative curvature), while the off-diagonal coupling K_{DG} prevents marginal directions in the generator from becoming neutrally stable; together, these ensure contraction toward equilibrium.

If K_{DD} or K_{DG} or both are rank-deficient, the corresponding zero-eigenvalue directions form a local subspace of equivalent equilibria, and the projected Jacobian is Hurwitz [1]. A Lyapunov function based on squared distance to the equilibrium set decreases along trajectories, by a LaSalle-type invariance argument, solutions converge to the equilibrium manifold asymptotically, and exponentially along the transverse directions [1].

Under Assumptions I–IV, the Jacobian at $\theta^\star$ is given as [1],

$$J = \begin{bmatrix} 2f''(0)\, K_{DD} & f'(0)\, K_{DG} \\ -\, f'(0)\, K_{DG}^\top & 0 \end{bmatrix}$$

If $Q = Q^\top \succ 0$ and $J^\top P + PJ = -Q$ for $P = P^\top \succ 0$, then with $V(\delta) = \delta^\top P \delta$,

$$\dot{V}(\delta) = \delta^\top (J^\top P + PJ)\delta = -\delta^\top Q \delta \leq -\lambda_{\min}(Q)\|\delta\|^2 \quad (10)$$

Thus, V strictly decreases and the origin is *exponentially stable*. Hurwitzness of J follows the block criterion with $Q' = -2f''(0)K_{DD} \succ 0$ and $P' = f'(0)K_{DG}$. If rank deficiencies correspond exactly to an equilibrium manifold, the same Lyapunov construction on the orthogonal complement and LaSalle's invariance yield convergence to that set [1]. Thus, we observe that the data generation process using GAN converges.

2) Convergence of FL: Each node k performs local updates for η_f local epochs, and

$$M_k^{e+1} = M_k^e - \eta_f \nabla_f M_k^e \quad (11)$$

where e denotes an epoch. After η_f epochs, each node sends its model update to the server. After receiving local model updates, the aggregated global model update at the server at round r is given as:

$$M^{r+1} = \frac{1}{K_r} \sum_{k=1}^{K_r} M_k^r \quad (12)$$

In [2], the authors demonstrated the convergence of FL and stated Theorem 2 as follows:

Theorem 2: Under the Lipschitz continuity, bounded variance, unbiased gradients, smoothness, and strong convexity, the FL-based framework converges to a stationary point of $\mathcal{L}(M_f)$.

Assuming that the global loss function $\mathcal{L}(M_f)$ is smooth, its change is expressed as [2]:

$$\begin{aligned} \mathcal{L}(M_f^{r+1}) \leq \mathcal{L}(M_f^r) + \nabla \mathcal{L}(M_f^r)^{\mathcal{R}} (M_f^{r+1} - M_f^r) \\ + \frac{C_1}{2} \|M_f^{r+1} - M_f^r\|^2 \end{aligned} \quad (13)$$

where C_1 is a constant and $C_1^2 > 0$.

Considering the expectations, we obtain the following equation [2]:

$$\mathbb{E}[\mathcal{L}(M_f^{r+1})] \leq \mathcal{L}(M_f^r) - \frac{\eta_f}{K_r} \|\nabla \mathcal{L}(M_f^r)\|^2 + \frac{C_2 \eta_f^2 C_1^2}{2K_r} \quad (14)$$

where C_2 is a constant and $C_2 > 0$.

Summing the inequality over $\mathcal{R}$ rounds and rearranging the terms, we obtain the following equation [2]:

$$\frac{1}{\mathcal{R}} \sum_{r=1}^{\mathcal{R}} \mathbb{E}[\|\nabla_f \mathcal{L}(M_f^r)\|^2] \leq \frac{\mathcal{L}(M_f^1) - \mathcal{L}(M_f^{\mathcal{R}})}{\eta_f \mathcal{R}} + \frac{C_2 \eta C_1^2}{2K_r} \quad (15)$$

As $\mathcal{R} \to \infty$, the term $\frac{\mathcal{L}(M_f^1) - \mathcal{L}(M_f^{\mathcal{R}+1})}{\eta \mathcal{R}}$ approaches zero, ensuring the following equation:

$$\lim_{\mathcal{R} \to \infty} \frac{1}{\mathcal{R}} \sum_{r=1}^{\mathcal{R}} \mathbb{E}[\|\nabla \mathcal{L}(M_f^r)\|^2] = 0 \quad (16)$$

This demonstrates that the global model update M_f converges to a stationary point of $\mathcal{L}(M_f)$.

REFERENCES

[1] V. Nagarajan and J. Z. Kolter, "Gradient descent gan optimization is locally stable," *Advances in neural information processing systems*, vol. 30, 2017.

[2] A. Mukherjee and R. Buyya, "Federated learning architectures: A performance evaluation with crop yield prediction application," *Software: Practice and Experience*, vol. 55, pp. 1165–1184, 2025.