Neumann Series Preconditioning for KKT Systems

Consider the indefinite KKT system $K_k z = r$ arising in nonlinear model predictive control, where $K_k = \begin{bmatrix} H_k & A_k^T \\ A_k & 0 \end{bmatrix}$ with H_k the Hessian and A_k the constraint Jacobian. We approximate K_k^{-1} via a truncated Neumann series using a simpler matrix D_k (typically block-diagonal):

$$M_k^{-1} = \sum_{j=0}^{p} (I - D_k^{-1} K_k)^j D_k^{-1}$$

The critical requirement is $\delta = \|I - D_k^{-1} K_k\| < 1$ for series convergence (spectral norm).

Series convergence. Assume $\delta < 1$ and let $E = I - D_k^{-1} K_k$. Since $D_k^{-1} K_k = I - E$, we have:

$$\begin{aligned} K_k^{-1} &= (D_k(I - E))^{-1} = (I - E)^{-1} D_k^{-1} && (1) \\ &= \left(\sum_{j=0}^{\infty} E^j \right) D_k^{-1} = \sum_{j=0}^{\infty} (I - D_k^{-1} K_k)^j D_k^{-1} && (2) \end{aligned}$$

The series converges because $\|E\| = \delta < 1$ ensures $\sum_{j=0}^{\infty} \|E\|^j = \frac{1}{1-\delta} < \infty$.

Truncation error. The error from using only $p + 1$ terms is:

$$\begin{aligned} \|K_k^{-1} - M_k^{-1}\| &= \left\| \sum_{j=p+1}^{\infty} E^j D_k^{-1} \right\| \leq \sum_{j=p+1}^{\infty} \|E\|^j \|D_k^{-1}\| && (3) \\ &= \|D_k^{-1}\| \cdot \frac{\delta^{p+1}}{1 - \delta} && (4) \end{aligned}$$

Thus the error in p greatly reduces when $\delta < 1$.

Eigenvalue analysis. Let λ_0 be an eigenvalue of $D_k^{-1} K_k$, so $(I - D_k^{-1} K_k)$ has eigenvalue $(1 - \lambda_0)$. The polynomial preconditioner computes:

$$p(\lambda_0) = \sum_{j=0}^{p} (1 - \lambda_0)^j = \frac{1 - (1 - \lambda_0)^{p+1}}{\lambda_0}$$

For $|1 - \lambda_0| \leq \delta < 1$, this approximates $1/\lambda_0$ with error:

$$\left| p(\lambda_0) - \frac{1}{\lambda_0} \right| = \frac{|1 - \lambda_0|^{p+1}}{|\lambda_0|} \leq \frac{\delta^{p+1}}{1 - \delta}$$

The preconditioned system $M_k^{-1} K_k$ has eigenvalues $\lambda = \lambda_0 \cdot p(\lambda_0) = 1 - (1 - \lambda_0)^{p+1}$, clustered near 1 with $|\lambda - 1| = O(\delta^{p+1})$. This yields condition number $\kappa(M_k^{-1} K_k) \approx \frac{1+\delta^{p+1}}{1-\delta^{p+1}}$, approaching 1 as p increases.

Iterative solver convergence. For conjugate gradient or MINRES applied to the preconditioned system, the error after l iterations satisfies:

$$\frac{\|x^{(l)} - x^*\|}{\|x^{(0)} - x^*\|} \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^l$$

With $\kappa \approx 1 + O(\delta^{p+1})$, the number of iterations for it to converge reduces dramatically.

Practical considerations. The method's viability depends on whether $\delta < 1$ is satisfied, which only applies for near-diagonal systems. For general indefinite KKT systems, obtaining $\delta < 1$ often requires D_k so close to K_k that computing D_k^{-1} becomes as expensive as solving the original system. This makes it largely impractical for single-system solves.

Time-varying systems. Should we be solving a time series of KKT systems, like $K_0, K_1, K_2, \ldots$ at successive time steps where each K_{k+1} differs only slightly from K_k, then there may be some benefit in reusing the preconditioner as long as $\|K_{k+j} - K_k\|$ remains small. Specifically, if the initial preconditioner satisfies $\delta_k = \|I - D_k^{-1} K_k\| < 1$ and there is a high enough sampling rate where the changes are sufficiently gradual, then this method could be much more viable compared to using it for single-system solves.