1 Background

Let $X = (X_t)_{t=1}^T$ be a zero-mean real vector-valued stochastic process with $X_t \in \mathbb{R}^n$ and $\mathbb{E}[X_t] = 0$ for $t = 1, 2, \ldots, T$. Consider the variance of the sum

$$\operatorname{Var}\left(\sum_{t=1}^T X_t\right) = \sum_{t=1}^T \sum_{t'=1}^T \operatorname{Cov}(X_t, X_{t'}) \tag{1}$$

$$= \sum_t \operatorname{Var}(X_t) + \sum_t \sum_{t'>t} \operatorname{Cov}(X_t, X_{t'}) + \operatorname{Cov}(X_{t'}, X_t), \tag{2}$$

where $\operatorname{Cov}(X_t, X_{t'}) = \mathbb{E}[X_t X_{t'}^\intercal]$ and $\operatorname{Var}(X_t) = \operatorname{Cov}(X_t, X_t)$. Given the marginal variances $\operatorname{Var}(X_t)$ and pairwise covariances $\operatorname{Cov}(X_t, X_{t'})$ of the elements of X, the variance above can be computed directly in time proportional to T^2. However, given some assumptions on X we might compute the variance (1) more efficiently. Below we list two possible assumptions.

X is stationary We say X is stationary if $\operatorname{Var}(X_t)$ is independent of t and $\operatorname{Cov}(X_t, X_{t'})$ depends only on the indices t and t' only through the difference $|t - t'|$ (noting that we always have $\operatorname{Cov}(X_t, X_{t'}) = \operatorname{Cov}(X_{t'}, X_t)$). In this case, we can simplify the expression in (1) as

$$\operatorname{Var}\left(\sum_{t=1}^T X_t\right) = T \operatorname{Var}(X_1) + \sum_{d=1}^{T-1} \operatorname{Cov}(X_1, X_{1+d}) + \operatorname{Cov}(X_{1+d}, X_1). \tag{3}$$

Given the autocovariance function of X, this expression can be computed in time linear in T.

X is a stationary linear autoregressive process An even stronger assumption is that X is not only stationary but also forms a Markov chain, and in addition that Markov chain takes the form of a linear autoregressive process:

$$X_{t+1} = AX_t + \epsilon_t, \tag{4}$$

where the spectral radius of A is less than 1 and where the noise process $\epsilon = (\epsilon_t)_{t=1}^{T-1}$ stationary and uncorrelated, with $\operatorname{Var}(\epsilon_t)$ independent of t and $\operatorname{Cov}(\epsilon_t, \epsilon_{t'}) = 0$ when $t \neq t'$. In this case, we can write a sum of consecutive covariances as a finite geometric series,

$$\sum_{d=1}^{T-1} \operatorname{Cov}(X_{1+d}, X_1) = \sum_{d=1}^{T-1} A^d \operatorname{Var}(X_1) = M \operatorname{Var}(X_1), \tag{5}$$

where $M \triangleq A(I - A)^{-1}(I - A^{T-1})$. Thus we can simplify the variance (1) as

$$\operatorname{Var}\left(\sum_{t=1}^T X_t\right) = T \operatorname{Var}(X_1) + M \operatorname{Var}(X_1) + \operatorname{Var}(X_1) M^\intercal. \tag{6}$$

Given A this expression can be computed in time constant in T.

Note that if X were a nonstationary linear autoregressive process, with

$$X_{t+1} = A_t X_t + \epsilon_t, \tag{7}$$

then to compute $\sum_{t,t'} \mathrm{Cov}(X_t, X_{t'})$ we would need to compute $\sum_{t,t'} \prod_{k=\min(t,t')}^{\max(t,t')} A_k$, which would require time quadratic in T.