Submitted to the Annals of Statistics

GEODESIC MARKOV CHAINS ON COVARIANCE MATRICES

By Makram Talih*,

City University of New York, Hunter College

This article is motivated by the problem of estimating the covariance structure in the multivariate Normal model when the underlying precision matrix is driven by a hidden Markov chain. We construct a Markov chain on the cone of positive definite matrices such that the path between consecutive matrices is a geodesic segment. The result is a geometric framework for multivariate stochastic volatility that includes both the matrix Beta-variate multiplicative evolution model of [24] and the matrix exponential GARCH model of [11].

1. Introduction. This paper is motivated by the problem of estimating the covariance structure in the multivariate Normal model when the underlying precision matrix is driven by a hidden Markov chain. We introduce a new class of dynamic models for precision matrices based on the intrinsic geometry of the cone $\mathcal{P}$ of real positive-definite matrices. When $\mathcal{P}$ is

*This work was completed while the author was a *New Researcher Fellow* at the Statistical and Applied Mathematical Sciences Institute, in connection with the Geometric Methods working group during the Fall 2006 program on High Dimensional Inference and Random Matrices. This work was supported by the National Science Foundation (NSF) under Agreement No. DMS-0112069. Any findings, conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF.

AMS 2000 subject classifications: 62H12, 53C22, 15A48, 62F15, 53C35, 15A52

Keywords and phrases: Dynamic Model, Stochastic Covariance, Matrix Exponential, Differential Geometry

imsart-aos ver. 2006/10/13 file: geodesic-paper2-AS.tex date: March 23, 2007

endowed with a Riemannian metric, an intrinsic distance can be defined and, as discussed in [21], corresponds to the Fisher information distance in the multivariate Normal model. The Markov chains we propose in this paper are such that the path between consecutive precision matrices is a geodesic (shortest intrinsic distance) segment in the cone $\mathcal{P}$, ensuring it remains well within $\mathcal{P}$ at all times. The main appeal of our use of the intrinsic geometry of the cone $\mathcal{P}$ is that, unlike in the usual Euclidean framework, nonpositive-definite symmetric matrices remain at an infinite (intrinsic) distance from the cone $\mathcal{P}$.

In the context of financial models for asset pricing and allocation, multivariate stochastic volatility models are well studied, cf. [5] for a recent review of the literature. [18] have contributed a general model for multivariate stochastic volatility in which the variance-covariance structure of asset returns follows a Wishart process, extending the traditional Normal-Wishart representation in the fixed-time case. For a positive-definite parameter matrix J, which the authors interpret as a measure of intertemporal sensitivity, a scalar parameter β, with $|\beta| < 1$, measuring persistence, and degrees of freedom $\alpha > p+1$, [18] assume a p-dimensional multivariate Normal distribution $N_p(0, \Sigma_{t+1})$ on (log-) returns at time $t+1$, together with a time-varying precision matrix $K_{t+1} = \Sigma_{t+1}^{-1}$ given by a Wishart distribution with degrees of freedom α and scale matrix $J^{1/2} K_t^{\beta} J^{1/2}/\alpha$. Thus, the conditional expectation of K_{t+1} given K_t is:

$$\mathbb{E}[K_{t+1} \mid K_t, J, \beta] = J^{1/2} K_t^{\beta} J^{1/2}. \tag{1.1}$$

The authors discuss how the interplay between the parameters β and J ensures stability.

The model of [18] stems from a long tradition of multiplicative updating of the covariance structure in dynamic matrix-evolution models, known in the Bayesian forecasting community as variance/covariance discounting. Variance-matrix discounting has proved extremely fruitful in empirical finance applications such as portfolio construction, and is a method that takes roots in the work of [20] and others, cf. [26], [1] and references therein. The idea of variance-matrix discounting is also found in the matrix Beta-variate multiplicative evolution model of [24]. [24] considered a vector auto-regression model with multivariate Normal errors and time-varying error precision matrices K_{t+1} at time $t+1$ given by

$$K_{t+1} = U_t^T \Theta_{t+1} U_t \,/\gamma, \tag{1.2}$$

where γ is a positive scalar, $K_t = U_t^T U_t$ is the Cholesky decomposition of the precision matrix K_t at time t with U_t upper-triangular, and Θ_{t+1} is a matrix of (unobserved) discount factors that, independently, follows a multivariate Beta type I distribution. – Just like its scalar counterpart, which is the distribution of the ratio $w_1/(w_1 + w_2)$ for independent Gamma random variables $w_i \sim \text{Gamma}(\alpha_i, 1)$, $i = 1, 2$, the Beta type I is the distribution of the ratio $W_1^{1/2}(W_1 + W_2)^{-1}W_1^{1/2}$ for independent Wishart matrices $W_i \sim \text{Wishart}_p(\alpha_i, I)$ with degrees of freedom α_i, $i = 1, 2$, and a common scale matrix I, the identity matrix, cf. [16].

In this paper, we introduce two new classes of dynamic models for precision matrices based on the intrinsic geometry of the cone $\mathcal{P}$ of real positive-definite matrices. One such class of models is described in section 3. We propose the following evolution equation:

$$K_{t+1} = K_t^{1/2}\mathrm{Exp}(\delta\tau_{t+1}B_{t+1})\ K_t^{1/2}, \tag{1.3}$$

where Exp denotes the matrix exponential, δ is a scalar with $0 < \delta < 1$, the τ_{t+1} are independent and identically distributed exponential random variables with mean one, independent of the K_t, and the B_{t+1} are independent and identically distributed symmetric random matrices, independent of the τ_{t+1} and the K_t, with norm $\|B_{t+1}\| \leq 1$. In section 3, we show that, when $\|B_{t+1}\| = 1$, $\delta\tau_{t+1}$ is the intrinsic distance along the geodesic segment between K_t and K_{t+1}. The latter is seen as a random segment on the geodesic curve originating from K_t with random initial velocity $K_t^{1/2}B_{t+1}K_t^{1/2}$. Our model (1.3) yields the following matrix of average "discounting factors":

$$\mathbb{E}[K_t^{-1/2}K_{t+1}K_t^{-1/2}] = \mathbb{E}[(I - \delta B_{t+1})^{-1}].$$

This is comparable to, yet more flexible than, the one obtained from (1.2) in [24]:

$$\mathbb{E}[(U_t^T)^{-1}K_{t+1}U_t^{-1}] = \mathbb{E}[\Theta_{t+1}/\gamma].$$

The added flexibility in our approach comes from the fact that the discounting factors in (1.3) are controlled by the matrices B_{t+1}. The latter are random matrices that lie inside the unit ball in $\mathcal{S}$, but that, otherwise, have no positive definiteness constraints. Both random matrix theory [8] and directional statistics [25] provide a rich collection of distributions for such matrices. In fact, in theorem 3.1, we show that, for a particular choice of matrices B_{t+1} – namely log-Beta type I distributed random matrices – model (1.2) of [24] is a special case of our model (1.3).

In section 5, we introduce an additional class of models based on the log-Euclidean geometry developed by [3] for the cone $\mathcal{P}$ in the context of diffusion tensor magnetic resonance imaging. We describe the evolution of precision matrices by a geodesic random walk on their logarithms:

$$\mathrm{Ln}(K_{t+1}) = \mathrm{Ln}(K_t) + \delta\tau_{t+1}\, \partial_{A_{t+1}}\mathrm{Ln}(K_t), \tag{1.4}$$

where $\partial_{A_{t+1}}\mathrm{Ln}(K_t)$ denotes the directional derivative of the matrix logarithm Ln at K_t in the direction of A_{t+1}, the A_{t+1} are symmetric random matrices dependent on K_t, $0 < \delta < 1$, and the τ_{t+1} are as before. We show that, when the directional derivative in (1.4) is appropriately standardized, the intrinsic distance along the geodesic segment between K_t and K_{t+1} remains $\delta\tau_{t+1}$, and we have, further, that

$$\mathbb{E}[\mathrm{Ln}(K_{t+1}) - \mathrm{Ln}(K_t)] = \delta\mathbb{E}[\partial_{A_{t+1}}\mathrm{Ln}(K_t)].$$

Thus, when the directions A_{t+1} are symmetric in distribution, the right-hand side in the above becomes zero, therefore the left-hand side becomes the expectation of a matrix-valued martingale difference, cf. section 5. Hence, the log-Euclidean model (1.4) generalizes the matrix exponential GARCH models of [11].

The Markov chains we propose remain well within the cone of positive definite matrices. Thus, other than for models of multivariate stochastic volatility, they can serve for Monte

Carlo inference from static data, and are easily adapted to sequential Monte Carlo.

The matrix exponential and logarithm have been used in [12] for approximate finite-sample and Bayesian inference techniques for the Normal covariance matrix, as well as in [6] to study dependence on explanatory variables of the covariance structure in a multivariate linear model.

The differential geometry associated with the Fisher information metric in the multivariate Normal model is well known, due to the seminal work of [2]. In the study of magnetic resonance imaging data, [7] have introduced a principal component analysis whose modes of variation are geodesic lines in the cone of diffusion tensors (i.e. 3×3 variance-covariance matrices). In the context of array signal processing, [22] make use of geodesics on the cone $\mathcal{P}$ to track time-varying linear subspaces. In computer vision, [19] build an algorithm for covariance tracking that is also based on intrinsic distance and its associated geometry.

The plan of the paper is as follows. Section 2 contains background material on Riemannian geometry. In section 3, we introduce a new class of dynamic models for precision matrices based on the Riemannian geometry of the cone $\mathcal{P}$ of positive-definite matrices. Using a convenient parameterization of geodesics on $\mathcal{P}$ (lemma 2.1), we show how our geodesic Markov chain model generalizes the matrix Beta-variate multiplicative evolution model of [24], cf. theorem 3.1. In section 5, we propose another class of models based on the log-Euclidean geometry of $\mathcal{P}$, which we review in section 4. The latter models produce matrix-logarithmic random walks that are comparable to the exponential GARCH model of [11]. The innovations in the log-Euclidean framework take a particularly simple form for the cone $\mathcal{P}$ (lemma 4.1), which makes them readily comparable to those in section 3. We discuss both classes of models in section 6 and, in theorem 6.2, derive a long-run distribution in the log-Euclidean case. Section 7 concludes.

2. Riemannian geometry of positive definite matrices.

The material in this section follows the exposition in [15]. Chapter 1 of [10] contains much of the foundational material on differentiable manifolds.

Let $\mathcal{S}(p)$ denote the space of real symmetric $p \times p$ matrices. In what follows, since the dimension p remains fixed, we write $\mathcal{S}$ for $\mathcal{S}(p)$. The Euclidean inner product on $\mathcal{S}$ is the Frobenius inner product

$$\langle A, B\rangle = \operatorname{tr}(A^T B) \overset{\text{here}}{=} \operatorname{tr}(AB).$$

The latter yields a matrix norm, called the Frobenius norm, which, for $B = (b_{ij}) \in \mathcal{S}$, is:

$$\|B\| = \langle B, B\rangle^{1/2} = \left(\sum_{i=1}^{p}\sum_{j=1}^{p} b_{ij}^2\right)^{1/2} = \left(\sum_{i=1}^{p} \lambda_i^2\right)^{1/2},$$

where the λ_i are the eigenvalues of B.

Let $\mathcal{P} = \{P \in \mathcal{S} : P > 0\}$ denote the set of positive definite matrices in $\mathcal{S}$. $\mathcal{P}$ is a convex open cone, in that $P + sQ \in \mathcal{P}$ for all $P, Q \in \mathcal{P}$ and $s \geq 0$. For each $P \in \mathcal{P}$, the *tangent space* $\mathcal{T}_P$ is defined as the collection of tangent "vectors" to $\mathcal{P}$ at P. It transpires that the tangent space $\mathcal{T}_I$ to $\mathcal{P}$ at the identity I is equal to the space of symmetric matrices $\mathcal{S}$, cf. [15]. Hence, for each $P \in \mathcal{P}$, the tangent space $\mathcal{T}_P$ is identified with $\mathcal{S}$ via the map $A \mapsto P^{-1}A$. A particular inner product on $\mathcal{T}_P$ is:

$$\langle A, B\rangle_P \overset{\text{def}}{=} \langle P^{-1}A, P^{-1}B\rangle_I = \langle P^{-1}A, P^{-1}B\rangle.$$

It induces a norm on $\mathcal{T}_P$ given by

$$\|A\|_P \overset{\text{def}}{=} \|P^{-1}A\| = \langle P^{-1}A, P^{-1}A\rangle^{1/2}.$$

The map $P \mapsto P^{-1}A$ is smooth. Therefore, the map $P \mapsto \langle\cdot,\cdot\rangle_P$ is smooth; the latter is called a Riemannian metric. Equipped with a Riemannian metric, $\mathcal{P}$ is a *Riemannian manifold*.

2.1. *Geodesics.* The *length* of a smooth (i.e. class C^∞) curve $\Gamma : [0,1] \to \mathcal{P}$ is defined as

$$\mathcal{L}(\Gamma) = \int_0^1 \|\Gamma'(s)\|_{\Gamma(s)}\, ds,$$

and the *energy* of Γ as

$$\mathcal{E}(\Gamma) = \frac{1}{2}\int_0^1 \|\Gamma'(s)\|^2_{\Gamma(s)}\, ds,$$

where $\Gamma'(s)$ denotes the derivative $d\Gamma(s)/ds$. The *intrinsic distance* between P and Q in the Riemannian manifold $\mathcal{P}$ is defined as

$$d(P,Q) = \inf\{\mathcal{L}(\Gamma) \mid \Gamma : [0,1] \to \mathcal{P}, \Gamma(0) = P, \Gamma(1) = Q\}, \tag{2.1}$$

where the infimum is taken over smooth curves Γ.

To paraphrase definition 1.4.2 in [10], to which the interested reader is referred to for a more detailed statement, we say that a smooth curve $\Gamma : [0,1] \to \mathcal{P}$ is a *geodesic* if it is a critical point of the energy functional $\mathcal{E}(\Gamma)$. This means that a geodesic curve Γ solves the so-called Euler-Lagrange ordinary differential equations for the energy $\mathcal{E}$ in the local coordinates. Loosely speaking, the "acceleration" along Γ remains zero. It transpires that length-minimizing smooth curves are geodesics, thus the infimum in (2.1) is achieved by geodesic curves.

For $\Gamma(0) = I \in \mathcal{P}$, the geodesic curve with initial velocity $\Gamma'(0) = B \in \mathcal{T}_I$ is given by

$$\Gamma(s) = \mathrm{Exp}(sB),$$

cf. [15], where the matrix exponential map $\mathcal{T}_I \to \mathcal{P}$ is everywhere defined on $\mathcal{T}_I$ via the formal power series expansion for the matrix exponential function:

$$\mathrm{Exp}(sB) = I + sB + \frac{s^2}{2!}B^2 + \cdots .$$

We group, here, some basic properties of the matrix exponential for later reference.

PROPOSITION 2.1 (Properties of the matrix exponential). *Let $B \in \mathcal{S}$.*

1. *If* $B = \mathrm{diag}(b_1, \cdots, b_p)$ *is a diagonal matrix, then* $\mathrm{Exp}(B) = \mathrm{diag}(\exp(b_1), \cdots, \exp(b_p))$;
2. *If W is an invertible matrix, then* $\mathrm{Exp}(WBW^{-1}) = W\mathrm{Exp}(B)W^{-1}$;
3. *In particular, for the spectral decomposition $B = VDV^T$ of B with D diagonal and V an orthogonal matrix,* $\mathrm{Exp}(B) = V\mathrm{Exp}(D)V^T$;
4. *For the determinant, we have* $\det(\mathrm{Exp}(B)) = \exp(\mathrm{tr}(B))$.

For an arbitrary $P \in \mathcal{P}$, the geodesic with $\Gamma(0) = P$ and $\Gamma'(0) = A \in \mathcal{T}_P$ is given by [15] as

$$\Gamma(s) = P^{1/2}\mathrm{Exp}(sP^{-1/2}AP^{-1/2})P^{1/2}.$$

Note that the existence and uniqueness of the geodesic $\Gamma : [0,1] \to \mathcal{P}$ with $\Gamma(0) = P \in \mathcal{P}$ and $\Gamma'(0) = A \in \mathcal{T}_P$ is, in general, a consequence of an existence and uniqueness theorem for the aforementionned Euler-Lagrange ordinary differential equations, cf. theorem 1.4.2 in [10]. On the other hand, the Hopf-Rinow theorem ([10], theorem 1.4.8) implies that $\mathcal{P}$ is geodesically complete. This means that the interval $[0,1]$ extends to $(-\infty, +\infty)$ and that, for any given pair $P, Q \in \mathcal{P}$, we can find a geodesic curve Γ such that $\Gamma(0) = P$ and $\Gamma(1) = Q$, namely by taking the initial velocity as $\Gamma'(0) = P^{1/2}BP^{1/2}$, for $B \in \mathcal{S}$ given by $B = \mathrm{Ln}(P^{-1/2}QP^{-1/2})$. Here, the (principal) matrix logarithm $\mathrm{Ln} : \mathcal{P} \to \mathcal{S}$ is defined by $\mathrm{Ln}(\mathrm{Exp}(B)) = B$ for $B \in \mathcal{S}$.

Henceforth, given the geodesic with $\Gamma(0) = P \in \mathcal{P}$ and $\Gamma'(0) = A \in \mathcal{T}_P$, we represent the initial velocity A using the symmetric matrix $B = P^{-1/2}AP^{-1/2}$, whence $\|B\| = \|A\|_P$ and

$$\Gamma(s) = P^{1/2}\mathrm{Exp}(sB)P^{1/2}.$$

2.2. *Intrinsic distance $d(P,Q)$.* Let the geodesic curve Γ with $\Gamma(0) = P \in \mathcal{P}$ and $\Gamma'(0) = \mathrm{Ln}(P^{-1/2}QP^{-1/2}) = B \in \mathcal{S}$ achieve the infimum in (2.1). Then, the intrinsic distance $d(P,Q)$ can be computed explicitly:

$$d(P,Q) = \|\mathrm{Ln}(P^{-1/2}QP^{-1/2})\| = \left(\sum_{i=1}^{p} \ln(\lambda_i)^2\right)^{1/2},$$

for eigenvalues λ_i of $P^{-1/2}QP^{-1/2}$. This was first shown by [14] (theorem 1). Since the λ_i are also eigenvalues of $P^{-1}Q$, one can compute the distance $d(P,Q)$, in practice, without invoking the matrix square root $P^{-1/2}$.

We group here some additional properties of $d(P,Q)$, cf. [9].

PROPOSITION 2.2 (Properties of $d(P,Q)$). *The following are true:*

1. *$d(P,Q)$ is a bona fide metric on $\mathcal{P}$;*
2. *Linearity: $d(P,Q) = |s|\|B\|$ for $Q = P^{1/2}\mathrm{Exp}(sB)P^{1/2}$;*
3. *$d(P,Q) = \infty$ if $P \in \mathcal{P}$ and $Q \in \mathcal{S} \setminus \mathcal{P}$;*
4. *Affine invariance: $d(CPC^T, CQC^T) = d(P,Q)$ for $C \in GL(p)$;*
5. *Invariance to inversion: $d(P,Q) = d(P^{-1}, Q^{-1})$.*

Property 3 above is important. If one were to use $d(P,Q)^2$ as a loss function in estimating a covariance matrix, then matrices $Q \in \mathcal{S} \setminus \mathcal{P}$ would not enter in the estimation procedure. This is a property that the corresponding Euclidean loss function $\|P - Q\|^2$ does not share.

Property 5 implies that, for $K = \Sigma^{-1}$ in the multivariate Normal $N_p(0,\Sigma)$, we have that $d(K,I) = d(I,\Sigma)$. Thus, even though they lead to entirely different model interpretations, both the covariance matrix and its inverse are the same distance away from the identity matrix.

2.3. *Alternative parametrization of geodesics on the cone $\mathcal{P}$.* We close this section with an important alternative parametrization of geodesics on the cone of $\mathcal{P}$ of positive definite matrices, which will be key to proving theorem 3.1 in the next section.

LEMMA 2.1. *For $P \in \mathcal{P}$, let $P = U^T U$ be any decomposition of P with U invertible, and let $A \in \mathcal{T}_P$. The geodesic curve Γ with $\Gamma(0) = P$ and $\Gamma'(0) = A$ is given by*

$$\Gamma(s) = U^T \mathrm{Exp}(s(U^T)^{-1} A U^{-1}) U.$$

PROOF. Let $V = P^{-1/2} U^T$. V is an orthogonal matrix. By property 2 of proposition 2.1,

$$\begin{aligned} P^{1/2}\mathrm{Exp}(sP^{-1/2}AP^{-1/2})P^{1/2} &= P^{1/2}VV^T\mathrm{Exp}(sP^{-1/2}AP^{-1/2})VV^TP^{1/2} \\ &= P^{1/2}V\mathrm{Exp}(sV^TP^{-1/2}AP^{-1/2}V)V^TP^{1/2} \\ &= U^T\mathrm{Exp}(s(U^T)^{-1}AU^{-1})U. \end{aligned}$$

□

3. Hidden geodesic Markov chain model. For a Normally distributed random vector $X_t \sim N_p(0, \Sigma_t)$ of dimension p, the model we propose in this section is such that the precision matrix $K_t = \Sigma_t^{-1}$ of X_t follows a hidden matrix-valued Markov chain in $(\mathcal{P}, d)$ that moves along geodesics, ensuring it stays inside $\mathcal{P}$ at all times.

3.1. *Data generating mechanism.* Suppose a matrix-valued process on $\mathcal{P}$ has been constructed and is in position K_t at time t. As noted earlier, for an arbitrary matrix $B_{t+1} \in \mathcal{S}$, a corresponding matrix $A_{t+1} \in \mathcal{T}_{K_t}$ is given by $A_{t+1} = K_t^{1/2} B_{t+1} K_t^{1/2}$, whence $\|A_{t+1}\|_{K_t} =$

$\|B_{t+1}\|$. We assume the data is generated as follows:

$$
\begin{aligned}
&B_{t+1} \overset{\text{IID}}{\sim} \text{ some distribution on } \{B \in \mathcal{S} : \|B\| \le 1\}, \\
&\tau_{t+1} \overset{\text{IID}}{\sim} \text{ Exponential}(1), \ \text{ independent of the } B_{t+1}, \\
&K_{t+1} = K_t^{1/2}\text{Exp}(\delta\tau_{t+1}B_{t+1})K_t^{1/2}, \ \text{ for some } 0 < \delta < 1, \qquad (3.1)\\
&X_{t+1} \mid K_{t+1} \sim N_d(0, \Sigma_{t+1}), \ \text{ for } \Sigma_{t+1} = K_{t+1}^{-1}.
\end{aligned}
$$

In the above, the τ_{t+1} are independent and identically distributed exponential random variables with mean one, independent of the K_t, and the B_{t+1} are independent and identically distributed symmetric random matrices, independent of the τ_{t+1} and the K_t. From (3.1), it follows that the intrinsic distance along the geodesic segment between K_t and K_{t+1} is given by

$$d(K_t, K_{t+1}) = \delta\tau_{t+1}\|B_{t+1}\| \le \delta\tau_{t+1},$$

with equality when $\|B_{t+1}\| = 1$. The distance $d(K_t, K_{t+1})$ measures the magnitude of the discount factor or innovation $K_t^{-1/2}K_{t+1}K_t^{-1/2}$ relative to the identity, in that, by property 4 in proposition 2.2,

$$d(K_t^{-1/2}K_{t+1}K_t^{-1/2}, I) = d(K_t, K_{t+1}).$$

That the model (1.2) of [24] is a special case of (3.1) above is shown in the following

THEOREM 3.1. *Suppose the B_{t+1}, τ_{t+1}, K_{t+1} and δ are as specified previously. Suppose, furthermore, that the B_{t+1} have an orthogonally invariant distribution. Then, for the Cholesky* *decomposition* $K_t = U_t^T U_t$ *of the precision matrix* K_t *with* U_t *upper-triangular, we have*

$$K_{t+1} \overset{\mathcal{D}}{=} U_t^T \mathrm{Exp}(\delta\tau_{t+1} B_{t+1}) U_t.$$

In particular, by taking $\Theta_{t+1}/\gamma = \mathrm{Exp}(\delta\tau_{t+1} B_{t+1})$ *to be a matrix-variate Beta type-I, the model (1.2) becomes a special case of (3.1).*

PROOF. This is an immediate consequence of lemma 2.1. □

3.2. *Submartingale properties implied by* $B \overset{\mathcal{D}}{=} -B$. Since $\|B_{t+1}\| \le 1$ and $0 < \delta < 1$, we may take conditional expectation given K_t and B_{t+1} on both sides of (3.1) to obtain:

$$\mathbb{E}[K_{t+1} \mid K_t, B_{t+1}] = K_t^{1/2}(I - \delta B_{t+1})^{-1} K_t^{1/2}. \tag{3.2}$$

For a scalar $\delta \in (0,1)$, an exponential random variable τ with mean one and any random matrix $B \in \mathcal{S}$, independent of τ, with $\|B\| \le 1$ and $B \overset{\mathcal{D}}{=} -B$, let

$$R \overset{\text{def}}{=} \mathbb{E}[\mathrm{Exp}(\pm\delta\tau B)] = \mathbb{E}[(I \pm \delta B)^{-1}], \tag{3.3}$$

$$r \overset{\text{def}}{=} \mathbb{E}[\exp(\pm\delta\tau \mathrm{tr}(B))] = \mathbb{E}[(I \pm \delta \mathrm{tr}(B))^{-1}]. \tag{3.4}$$

Then, on the one hand, Jensen's inequality (in Loewner ordering) implies that the matrix $R-I$ is nonnegative definite. Since the B_{t+1} are IID, it follows from (3.2) that

$$\mathbb{E}[K_{t+1} - K_t \mid K_t] = K_t^{1/2}\,(R - I)\,K_t^{1/2},$$

which is nonnegative definite, meaning that the sequence of precision matrices $\{K_t\}$ is a matrix-valued submartingale. On the other hand, using property 4 in proposition 2.1, Jensen's inequality implies $r \geq 1$, hence the determinants $\det(K_t)$ form a submartingale sequence. By symmetry, both Σ_t and $\det(\Sigma_t)$ also form submartingales sequences.

4. Log-Euclidean geometry of positive definite matrices.

The material in this section is based on the log-Euclidean geometry developed by [3] for the cone $\mathcal{P}$ in the context of diffusion tensor magnetic resonance imaging. [3] define logarithmic multiplication by:

$$P \odot Q \stackrel{\text{def}}{=} \text{Exp}[\text{Ln}(P) + \text{Ln}(Q)].$$

Then, by proposition 3.2 in [3], $(\mathcal{P}, \odot)$ becomes a group, where the (now commutative) group multiplication generalizes ordinary matrix multiplication in that $P \odot Q = PQ$ when P and Q commute, and where the group inverse thus satisfies $P^{-1} \odot P = P^{-1}P = I$. Moreover, by introducing the log-scalar multiplication

$$s \circledast P \stackrel{\text{def}}{=} \text{Exp}[s\text{Ln}(P)] = P^s,$$

[3] make $\mathcal{P}$ a vector space in its own right. Further, theorem 3.3 in [3] shows that the map $(P, Q) \mapsto P^{-1} \odot Q$ remains smooth, and, hence, that $(\mathcal{P}, \odot)$ is a Lie group. Thus, the tangent space $\mathcal{T}_I$ to $\mathcal{P}$ at the identity element I remains the space of symmetric matrices $\mathcal{S}$, and an

inner product is defined on $\mathcal{T}_P$ by

$$\langle A, B\rangle_P = \langle \partial_A \mathrm{Ln}(P),\ \partial_B \mathrm{Ln}(P)\rangle = \mathrm{tr}(\partial_A \mathrm{Ln}(P)\ \partial_B \mathrm{Ln}(P)),$$

where $\partial_A \mathrm{Ln}(P)$ denotes the directional derivative in the direction of A of the Ln map evaluated at P, cf. [3] (corollary 3.9).

4.1. *Log-Euclidean geodesics.* By [3] (corollary 3.9), the log-Euclidean geodesic curve with $\Gamma(0) = P \in \mathcal{P}$ and $\Gamma'(0) = A \in \mathcal{T}_P$ is given by

$$\Gamma(s) = \mathrm{Exp}(\mathrm{Ln}(P) + s\partial_A \mathrm{Ln}(P)). \tag{4.1}$$

4.2. *Log-Euclidean intrinsic distance $d_{le}(P, Q)$.* In (4.1), for $Q = \Gamma(1)$, the resulting intrinsic log-Euclidean distance between P and Q in $\mathcal{P}$, which we denote by $d_{\mathrm{le}}(P, Q)$, is given simply by the Frobenius distance between their matrix logarithms, and it can be expressed in terms of the norm of the directional derivative of the matrix logarithm in the direction of A as follows:

$$d_{\mathrm{le}}(P, Q) = \|\mathrm{Ln}(P) - \mathrm{Ln}(Q)\| = \|\partial_A \mathrm{Ln}(P)\|.$$

PROPOSITION 4.1 (Properties of $d_{\mathrm{le}}(P, Q)$, cf. [3]). *The following are true:*

1. *$d_{le}(P, Q)$ is a bona fide metric on $\mathcal{P}$;*
2. *Linearity: $d_{le}(P, Q) = |s|\|\partial_A \mathrm{Ln}(P)\|$ for $Q = \mathrm{Exp}(\mathrm{Ln}(P) + s\partial_A \mathrm{Ln}(P))$;*
3. *$d_{le}(P, Q) = \infty$ if $P \in \mathcal{P}$ and $Q \in \mathcal{S} \setminus \mathcal{P}$;*

4. Scale invariance: $d_{le}(sP, sQ) = d_{le}(P, Q)$ *for scalars* $s > 0$*;*

5. Rotation invariance: $d_{le}(OPO^T, OQO^T) = d_{le}(P, Q)$ *for* $O \in SO(p)$*;*

6. Invariance to inversion: $d_{le}(P, Q) = d_{le}(P^{-1}, Q^{-1})$*.*

4.3. *Alternative parameterization of log-Euclidean geodesics in* $\mathcal{P}$*.* We introduce, here, an alternative parameterization of geodesics in the log-Euclidean framework that takes explicitly into account the form of directional derivatives $\partial_A \mathrm{Ln}(P)$ in (4.1). This will afford us a more transparent data-generating mechanism in section 5.

For $P \in \mathcal{P}$, consider the spectral decomposition $P = V^T D V$, where V is the orthogonal matrix of eigenvectors and $D = \mathrm{diag}(d_1, \cdots, d_p)$ is the diagonal matrix of (ordered) eigenvalues of P. For $A \in \mathcal{T}_P$, let $\Upsilon \in \mathcal{S}$ be given by $\Upsilon = VAV^T$. Theorem 4.8 of [17] shows that the directional derivative $\partial_A \mathrm{Ln}(P)$ satisfies

$$\partial_A \mathrm{Ln}(P) = V^T \partial_\Upsilon \mathrm{Ln}(D) V. \tag{4.2}$$

Further, using the so-called *Loewner matrix* $\Delta_D(\mathrm{Ln})$ whose entries are

$$\left(\Delta_D(\mathrm{Ln})\right)_{ij} = \begin{cases} \dfrac{\ln d_i - \ln d_j}{d_i - d_j}, \; d_i \neq d_j, \\ \dfrac{1}{d_i}, \; d_i = d_j, \end{cases}$$

the directional derivative $\partial_\Upsilon \mathrm{Ln}(D)$ in (4.2) is computed as the Hadamard product

$$\partial_\Upsilon \mathrm{Ln}(D) = \Upsilon \cdot \Delta_D(\mathrm{Ln}). \tag{4.3}$$

LEMMA 4.1 (Alternative parameterization of log-Euclidean geodesics in $\mathcal{P}$). *Let $P \in \mathcal{P}$, and consider its spectral decomposition $P = V^T D V$ given above. Let $B \in \mathcal{S}$ be an arbitrary matrix. Define $\Upsilon \in \mathcal{S}$ by*

$$\Upsilon = B \cdot \Delta_D^-(\mathrm{Ln}),$$

where the entries of $\Delta_D^-(\mathrm{Ln})$ are the reciprocals of the entries of the Loewner matrix $\Delta_D(\mathrm{Ln})$:

$$\left(\Delta_D^-(\mathrm{Ln})\right)_{ij} = \begin{cases} \dfrac{d_i - d_j}{\ln d_i - \ln d_j}, \ d_i \neq d_j, \\ \\ d_i, \ d_i = d_j. \end{cases}$$

Then $\Upsilon \cdot \Delta_D(\mathrm{Ln}) = B$ and, the log-Euclidean geodesic curve with $\Gamma(0) = V^T D V$ and $\Gamma'(0) = V^T \Upsilon V$ is given by

$$\Gamma(s) = V^T \mathrm{Exp}(\mathrm{Ln}(D) + sB) V.$$

PROOF. Let $A = V^T \Upsilon V$. $A \in \mathcal{T}_{\mathcal{P}}$. By (4.2) and (4.3), the directional derivative in (4.1) becomes $\partial_A \mathrm{Ln}(P) = V^T B V$. The result now follows from property 3 of proposition 2.1. □

5. Log-Euclidean hidden geodesic random walk. In this section, the evolution over time of a Normally distributed random vector $X_t \sim N_p(0, \Sigma_t)$ is such that the precision matrix $K_t = \Sigma_t^{-1}$ follows a hidden matrix-valued log-Euclidean random walk in $(\mathcal{P}, d_{\mathrm{le}})$ that moves along log-Euclidean geodesic segments.

5.1. *Data generating mechanism.* As before, let the spectral decomposition $K_t = V_t^T D_t V_t$ of the precision matrix $K_t \in \mathcal{P}$ at time t be given. Let $B_{t+1} \in \mathcal{S}$ be an arbitrary matrix. As in lemma 4.1, define $\Upsilon_{t+1} \in \mathcal{S}$ by

$$\Upsilon_{t+1} = B_{t+1} \cdot \Delta^{-}_{D_t}(\mathrm{Ln}),$$

where the entries of $\Delta^{-}_{D_t}(\mathrm{Ln})$ are the reciprocals of the entries of the Loewner matrix $\Delta_{D_t}(\mathrm{Ln})$.

Now, consider the following log-Euclidean data-generating mechanism:

$$B_{t+1} \overset{\mathrm{IID}}{\sim} \text{ some distribution on } \{B \in \mathcal{S} : \|B\| \le 1\},$$

$$\tau_{t+1} \overset{\mathrm{IID}}{\sim} \text{ Exponential}(1), \;\; \text{independent of the } B_{t+1},$$

$$\mathrm{Ln}(K_{t+1}) = \mathrm{Ln}(K_t) + \delta\tau_{t+1} V_t^T B_{t+1} V_t, \;\; \text{for } \; 0 < \delta < 1, \tag{5.1}$$

$$X_{t+1} \mid K_{t+1} \sim N_d(0, \Sigma_{t+1}), \;\; \text{for } \Sigma_{t+1} = K_{t+1}^{-1}.$$

In the above, the τ_{t+1} are independent and identically distributed exponential random variables with mean one, independent of the K_t, and the B_{t+1} are independent and identically distributed symmetric matrices, independent of the τ_{t+1} and the K_t. By lemma 4.1, the segment between K_t and K_{t+1} is a segment on the log-Euclidean geodesic curve starting at K_t with initial velocity $A_{t+1} = V_t^T \Upsilon_{t+1} V_t$. Further, the log-Euclidean distance along that segment is given by

$$d_{\mathrm{le}}(K_t, K_{t+1}) = \delta\tau_{t+1}\|B_{t+1}\| \le \delta\tau_{t+1},$$

with equality when $\|B_{t+1}\| = 1$. Recall that the distance $d_{\mathrm{le}}(K_t, K_{t+1}) = \|\mathrm{Ln}(K_{t+1}) - \mathrm{Ln}(K_t)\|$.

5.2. *Submartingale properties implied by $B \stackrel{\mathcal{D}}{=} -B$.* If $B \stackrel{\mathcal{D}}{=} -B$, then, from (5.1), we obtain

$$\mathbb{E}[\mathrm{Ln}(K_{t+1}) - \mathrm{Ln}(K_t) \mid K_t] = 0.$$

Thus, $\mathrm{Ln}(K_t)$ forms a matrix-valued martingale, of which the exponential GARCH model of [11] is a special case. Further, as $\|B_{t+1}\| \leq 1$ and $0 < \delta < 1$, we may take conditional expectation given K_t and B_{t+1} on both sides of (5.1) after exponentiating:

$$\mathbb{E}[K_{t+1} \mid K_t, B_{t+1}] = K_t^{1/2} \odot (I - \delta V_t^T B_{t+1} V_t)^{-1} \odot K_t^{1/2}.$$

Just as in section 3.2, but this time conditioning on K_t, define

$$R_t \stackrel{\text{def}}{=} \mathbb{E}[\mathrm{Exp}(\pm\delta\tau V_t^T B V_t) \mid K_t] = \mathbb{E}[(I \pm \delta V_t^T B V_t)^{-1} \mid K_t],$$

$$r_t \stackrel{\text{def}}{=} \mathbb{E}[\exp(\pm\delta\tau \mathrm{tr}(V_t^T B V_t)) \mid K_t] = \mathbb{E}[(I \pm \delta \mathrm{tr}(V_t^T B V_t))^{-1} \mid K_t], \tag{5.2}$$

where $0 < \delta < 1$ and $B \in \mathcal{S}$ with $B \stackrel{\mathcal{D}}{=} -B$. Note that $\mathrm{tr}(V_t^T B V_t) = \mathrm{tr}(B)$, since V_t is orthogonal. Hence, $r_t \equiv r$, with r as in (3.4). Note that if, in addition to being symmetric, the distribution of B is orthogonally invariant, then we will also have $R_t \equiv R$, with R as in (3.3). Now, Jensen's inequality implies that the matrix $R_t - I$ is nonnegative definite, which means the precision matrices K_t form a matrix-valued submartingale. On the other hand, the determinants $\det(K_t)$ form a submartingale sequence identical to the one obtained in section 3. By symmetry, both the covariance matrices Σ_t and their determinants are also submartingales.

5.3. *Extension to log-Euclidean autoregressive models.* The multivariate stochastic volatility model of [18], which we discussed in the introductory section of this article, is a Wishart random walk where, for $J = I$ – the identity matrix, – in (1.1), $\mathbb{E}[K_{t+1} \mid K_t] = K_t^{\beta}$. A comparable class of models can be obtained in the log-Euclidean framework of this section by replacing (5.1) in the data-generating mechanism with

$$\mathrm{Ln}(K_{t+1}) = \beta \mathrm{Ln}(K_t) + \delta \tau_{t+1} V_t^T B_{t+1} V_t,$$

for $|\beta| < 1$. In this model, the segment between K_t^{β} and K_{t+1} becomes a segment on the log-Euclidean geodesic curve starting at K_t^{β} with initial velocity

$$A_{t+1} = V_t^T (B_{t+1} \cdot \Delta^{-}_{D_t^{\beta}}(\mathrm{Ln})) V_t.$$

6. Comparative study. The log-Euclidean framework of section 5 has some advantages over the framework of section 3. First, we obtain *bona fide* random walks for the log-precision matrices. Second, since calculations are, in effect, done in the space $\mathcal{S}$, which is flat as it is a Euclidean space, numerical procedures are relatively more stable than when working directly in the (negatively) curved manifold $\mathcal{P}$. The drawback of the log-Euclidean framework is that, unlike the multiplicative innovations or discount factors of section 3, the log-Euclidean innovations are no longer independent of the present state K_t. Nonetheless, we show in theorem 6.1 that, for orthogonally invariant matrices B_{t+1}, the log-Euclidean innovations in (5.1) become distributionally equivalent to the multiplicative innovations of section 3.

THEOREM 6.1 (Innovations). *Let $B \in \mathcal{S}$ be a random matrix with an orthogonally invariant distribution. For $P \in \mathcal{P}$, let $P = V^T DV$ be the spectral decomposition of P, with V orthogonal and D diagonal. Consider the matrices Q_1 and Q_2 in $\mathcal{P}$ defined by*

$$Q_1 = P^{1/2}\mathrm{Exp}(B)P^{1/2} = \Gamma_1(1),$$

$$Q_2 = \mathrm{Exp}(\mathrm{Ln}(P) + V^T BV) = \Gamma_2(1).$$

In the above, Γ_1 is the geodesic curve in $(\mathcal{P}, d)$ with $\Gamma_1(0) = P$ and $\Gamma_1'(0) = P^{1/2}BP^{1/2}$, while Γ_2 is the geodesic curve in $(\mathcal{P}, d_{le})$ with $\Gamma_2(0) = P$ and $\Gamma_2'(0) = V^T(B \cdot \Delta_D^{-}\mathrm{Ln})V$. The innovations $P^{-1/2}Q_1P^{-1/2}$ and $P^{-1/2} \odot Q_2 \odot P^{-1/2}$ are equal in distribution.

PROOF. By logarithmic multiplication and scaling, we have

$$P^{-1/2} \odot Q_2 \odot P^{-1/2} = \mathrm{Exp}\big(\mathrm{Ln}(Q_2) - \mathrm{Ln}(P)\big) = \mathrm{Exp}(V^T BV).$$

Since the matrix B is orthogonally invariant, the right-hand side is equal in distribution to $\mathrm{Exp}(B)$, which is the innovation $P^{-1/2}Q_1P^{-1/2}$. □

THEOREM 6.2 (Limiting behavior). *For $t = 0, 1, \cdots$, let the spectral decomposition $K_t = V_t^T D_t V_t$ of the precision matrix $K_t \in \mathcal{P}$ at time t be given, where V_t is the orthogonal matrix of eigenvectors and D_t is the diagonal matrix of (ordered) eigenvalues of K_t. Let the τ_{t+1} be independent and identically distributed exponential random variables with mean one, independent of the K_t, and the B_{t+1} be independent and identically distributed symmetric matrices,* *independent of the τ_{t+1} and the K_t, with $\|B_{t+1}\| \le 1$. Suppose, moreover, that the distribution of the B_{t+1} is orthogonally invariant. Then, for the log-Euclidean random walk given by (5.1), we have that, as $t \to \infty$,*

$$(K_0^{-1/2} \odot K_t \odot K_0^{-1/2})^{1/t} \to \mathrm{Exp}(\delta \bar{B}),$$

where $\bar{B} = \mathbb{E}[B_1]$ and convergence is in probability.

PROOF. By (5.1), whenever the B_{t+1} have an orthogonally invariant distribution, we get

$$\mathrm{Ln}(K_{t+1}) - \mathrm{Ln}(K_t) \overset{\mathcal{D}}{=} \delta \tau_{t+1} B_{t+1},$$

which is independent of K_t. It follows that, for $t \ge 1$,

$$\mathrm{Ln}(K_t) - \mathrm{Ln}(K_0) \overset{\mathcal{D}}{=} \delta \sum_{s=0}^{t-1} \tau_{s+1} B_{s+1}.$$

The above sum is that of independent and identically distributed random matrices with mean $\mathbb{E}[\tau_{s+1} B_{s+1}] = \bar{B}$. By the law of large numbers,

$$\frac{1}{t} \sum_{s=0}^{t-1} \tau_{s+1} B_{s+1} \to \bar{B}$$

in probability as $t \to \infty$. Hence

$$\frac{1}{t}(\mathrm{Ln}(K_t) - \mathrm{Ln}(K_0)) \to \bar{B}$$

in probability as $t \to \infty$. The result now follows from logarithmic multiplication and scaling. $\square$

7. Discussion. The models we propose in this paper are hidden Markov chain models for the covariance structure in the multivariate Normal model. A Markov chain on the cone of real positive definite matrices is constructed in such a way that the path between consecutive matrices is a geodesic segment. The result is a geometric framework that includes as special cases both the matrix Beta-variate multiplicative evolution model of [24] and the matrix exponential GARCH model of [11]. Such Markov chains can also serve as Metropolis chains in a Markov chain Monte Carlo inference about the covariance or precision matrix in the multivariate Normal model, and they are easily adapted to online inference.

Determining the Riemannian geometry of the cone of positive definite matrices $\mathcal{P}_G$ with a pattern of zeroes dictated by the (absent) edges of an undirected graph G, such as arise in graphical models, remains an open problem. Those models provide parsimonious precision matrices and less volatile portfolio allocation weights in portfolio selection problems, cf. [23] and [4]. [13] (theorem 2.2) have shown that for a certain class of decomposable graphs called homogeneous graphs, the cone $\mathcal{P}_G$ is a homogeneous cone, i.e. its automorphism group acts on it transitively. Having a homogeneous cone facilitates the construction of the exponential map $\mathcal{P}_G \to \mathcal{T}_{\mathcal{P}_G}$ and the ensuing geodesic curves. The geodesic Markov chains proposed in this paper would provide a much needed tool for the study of multivariate stochastic volatility in the Gaussian graphical model.

References.

[1] Aguilar, O. and West, M. (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics* **18**, 3, 338–357.

[2] Amari, S. (1985). *Differential-Geometric Methods in Statistics*. Lecture Notes in Statistics, Vol. **28**. Springer-Verlag, New York, NY.

[3] Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. (2007). Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications* **29**, 1, 328–347.

[4] Carvalho, C. and West, M. (2007). Dynamic matrix variate graphical models. *Bayesian Analysis* **2**, 1, 69–98.

[5] Chib, S., Nardari, F., and Shephard, N. (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics 134*, 341–371.

[6] Chiu, T. Y. M., Leonard, T., and Tsui, K.-W. (1996). The matrix-logarithmic covariance model. *Journal of the American Statistical Association* **91**, 433, 198–210.

[7] Fletcher, P. T. and Joshi, S. (2004). Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Proceedings of the ECCV 2004 Workshop on Computer Vision Approaches to Medical Image Analysis*. Lecture Notes in Computer Science, Vol. **3117**. Springer-Verlag, New York, NY, 87–98.

[8] Forrester, P. (2005). Log-gases and random matrices. Book preprint available online at http://www.ms.unimelb.edu.au/∼matpjf/matpjf.html.

[9] Förstner, W. and Moonen, B. (1999). A Metric for covariance matrices. Technical report, Stuttgart University, Department of Geodesy and Geoinformatics.

[10] Jost, J. (2002). *Riemannian Geometry and Geometric Analysis, Third Edition*. Springer, Berlin.

[11] Kawakatsu, H. (2006). Matrix exponential GARCH. *Journal of Econometrics 134*, 95–128.

[12] Leonard, T. and Hsu, J. S. J. (1992). Bayesian inference for a covariance matrix. *Annals of Statistics* **20**, 4, 1669–1696.

[13] Letac, G. and Massam, H. (2007). Wishart distributions for decomposable graphs. *Annals of Statistics* **35**, 3. To appear.

[14] Maass, H. (1971). *Siegel's Modular Forms and Dirichlet Series.* Lecture Notes in Mathematics, Vol. **216**. Springer-Verlag, New York, NY.

[15] Moakher, M. (2005). A Differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications* **26**, 3, 735–747.

[16] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* John Wiley & Sons, New York, NY.

[17] Najfeld, I. and Havel, T. F. (1995). Derivatives of the matrix exponential and their computation. *Advances in Applied Mathematics* **16**, 3, 259–375.

[18] Philipov, A. and Glickman, M. E. (2006). Multivariate stochastic volatility via Wishart processes. *Journal of Business and Economic Statistics* **24**, 3, 313–328.

[19] Porikli, F., Tuzel, O., and Meer, P. (2006). Covariance tracking using model update based on Lie algebra. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

[20] Quintana, J. M. and West, M. (1987). An Analysis of international exchange rates using multivariate DLM's. *The Statistician* **36**, 2/3, 275–281.

[21] Smith, S. T. (2005). Covariance, subspace, and intrinsic Cramér-Rao bounds. *IEEE Transactions on Signal Processing* **53**, 5, 1610–1630.

[22] Srivastava, A. and Klassen, E. (2004). Bayesian and geometric subspace tracking. *Advances in Applied Probability* **36**, 1, 43–56.

[23] Talih, M. and Hengartner, N. (2005). Structural learning with time-varying components: Tracking the cross-section of financial time series. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* **67**, 3, 321–341.

[24] Uhlig, H. (1997). Bayesian vector autoregressions with stochastic volatility. *Econometrica* **65**, 1, 59–73.

[25] Ulrich, G. (1984). Computer generation of distributions on the m-sphere. *Applied Statistics* **33**, 2, 158–163.

[26] West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*, 2nd ed. Springer-Verlag, New York, NY.

City University of New York
Hunter College
Department of Mathematics and Statistics
695 Park Avenue, New York, NY, 10021
E-mail: makram.talih@hunter.cuny.edu