Glossary

**Adjoint** – For a finite-dimensional linear map (i.e., a matrix $A$), the adjoint $A^*$ is given by the complex conjugate transpose of the matrix. In the infinite-dimensional context, the adjoint $\mathcal{A}^*$ of a linear operator $\mathcal{A}$ is defined so that $\langle \mathcal{A} f, g \rangle = \langle f, \mathcal{A}^* g \rangle$, where $\langle \cdot, \cdot \rangle$ is an inner product.

**Closed-loop control** – A control architecture where the actuation is informed by sensor data about the output of the system.

**Coherent structure** – A spatial mode that is correlated with the data from a system.

**Compressed sensing** – The process of reconstructing a high-dimensional vector signal from a random undersampling of the data using the fact that the high-dimensional signal is sparse in a known transform basis, such as the Fourier basis.

**Compression** – The process of reducing the size of a high-dimensional vector or array by approximating it as a sparse vector in a transformed basis. For example, MP3 and JPEG compression use the Fourier basis or wavelet basis to compress audio or image signals.

**Control theory** – The framework for modifying a dynamical system to conform to desired engineering specifications through sensing and actuation.

**Data matrix** – A matrix where each column vector is a snapshot of the state of a system at a particular instant in time. These snapshots may be sequential in time, or they may come from an ensemble of initial conditions or experiments.

**DMD amplitude** – The amplitude of a given DMD mode as expressed in the data. These amplitudes may be interpreted as the significance of a given DMD mode, similar to the power spectrum in the FFT.

**DMD eigenvalues** – Eigenvalues of the best-fit DMD operator $A$ (see dynamic mode decomposition) representing an oscillation frequency and a growth or decay term.

**DMD mode (also dynamic mode)** – An eigenvector of the best-fit DMD operator $A$ (see dynamic mode decomposition). These modes are spatially coherent and oscillate in time at a fixed frequency and a growth or decay rate.

**Dynamic mode decomposition (DMD)** – The leading eigendecomposition of a best-
fit linear operator $A = X'X'$ that propagates the data matrix $X$ into a future data matrix $X'$. The eigenvectors of $A$ are DMD modes and the corresponding eigenvalues determine the time dynamics of these modes.

**Dynamical system** – A mathematical model for the dynamic evolution of a system. Typically, a dynamical system is formulated in terms of ordinary differential equations on a state-space. The resulting equations may be linear or nonlinear and may also include the effect of actuation inputs and represent outputs as sensor measurements of the state.

**Eigensystem realization algorithm (ERA)** – A system identification technique that produces balanced input-output models of a system from impulse response data. ERA has been shown to produce equivalent models to BPOD and DMD under some circumstances.

**Emission** – The measurement functions for an HMM.

**Equation-free modeling (EFM)** – The process of characterizing a dynamical system through measurement data as opposed to first principles modeling of physics. It is important to note that equation-free modeling does result in equations that model the system, but these methods are data driven and do not start with knowledge of the high-fidelity governing equations.

**Feedback control** – Closed-loop control where sensors measure the downstream effect of actuators, so that information is fed back to the actuators. Feedback is essential for robust control where model uncertainty and instability may be counteracted with fast sensor feedback.

**Feedforward control** – Control where sensors measure the upstream disturbances to a system, so that information is fed forward to actuators to cancel disturbances proactively.

**Hidden Markov model (HMM)** – A Markov model where there is a hidden state that is only observed through a set of measurements known as emissions.

**Hilbert space** – A generalized vector space with an inner product. When used in this text, a Hilbert space typically refers to an infinite-dimensional function space. These spaces are also complete metric spaces, providing a sufficient mathematical framework to enable calculus on functions.

**Incoherent measurements** – Measurements that have a small inner product with the basis vectors of a sparsifying transform. For instance, single-pixel measurements (i.e., spatial delta functions) are incoherent with respect to the spatial Fourier transform basis, since these single-pixel measurements excite all frequencies and do not preferentially align with any single frequency.

**Kalman filter** – An estimator that reconstructs the full state of a dynamical system from measurements of a time series of the sensor outputs and actuation inputs. A Kalman filter is itself a dynamical system that is constructed for observable systems to stably converge to the true state of the system. The Kalman filter is optimal for linear
systems with Gaussian process and measurement noise of a known magnitude.

**Koopman eigenfunction** – An eigenfunction of the Koopman operator. These eigenfunctions correspond to measurements on the state-space of a dynamical system that form intrinsic coordinates. In other words, these intrinsic measurements will evolve linearly in time despite the underlying system being nonlinear.

**Koopman operator** – An infinite-dimensional linear operator that propagates measurement functions from an infinite-dimensional Hilbert space through a dynamical system.

**Least-squares regression** – A regression technique where a best-fit line or vector is found by minimizing the sum of squares of the error between the model and the data.

**Linear system** – A system where superposition of any two inputs results in the superposition of the two corresponding outputs. In other words, doubling the input doubles the output. Linear time-invariant dynamical systems are characterized by linear operators, which are represented as matrices.

**Low rank** – A property of a matrix where the number of linearly independent rows and columns is small compared with the size of the matrix. Generally, low-rank approximations are sought for large data matrices.

**Markov model** – A probabilistic dynamical system where the state vector contains the probability that the system will be in a given state; thus, this state vector must always sum to unity. The dynamics are given by the Markov transition matrix, which is constructed so that each row sums to unity.

**Markov parameters** – The output measurements of a dynamical system in response to an impulsive input.

**Moore’s law** – The observation that transistor density, and hence processor speed, increase exponentially in time. Moore’s law is commonly used to predict computational power and the associated increase in the scale of problem that will be computationally feasible.

**Multiscale** – The property of having many scales in space and/or time. Many systems, such as turbulence, exhibit spatial and temporal scales that vary across many orders of magnitude.

**Observable function** – A function that measures some property of the state of a system. Observable functions are typically elements of a Hilbert space.

**Overdetermined system** – A system $Ax = b$ where there are more equations than unknowns. Usually there is no exact solution $x$ to an overdetermined system, unless the vector $b$ is in the column space of $A$.

**Perron–Frobenius operator** – The adjoint of the Koopman operator; an infinite-dimensional operator that advances probability density functions through a dynamical system.
**Power spectrum** – The squared magnitude of each coefficient of a Fourier transform of a signal. The power corresponds to the amount of each frequency required to reconstruct a given signal.

**Principal component** – A spatially correlated mode in a given data set, often computed using the SVD of the data after the mean has been subtracted.

**Principal component analysis (PCA)** – A decomposition of a data matrix into a hierarchy of principal component vectors that are ordered from most correlated to least correlated with the data. PCA is computed by taking the SVD of the data after subtracting the mean. In this case, each singular value represents the variance of the corresponding principal component (singular vector) in the data.

**Proper orthogonal decomposition (POD)** – The decomposition of data from a dynamical system into a hierarchical set of orthogonal modes, often using the SVD. When the data consists of velocity measurements of a system, such as an incompressible fluid, then POD orders modes in terms of the amount of energy these modes contain in the given data.

**Pseudoinverse** – The pseudoinverse generalizes the matrix inverse for nonsquare matrices, and is often used to compute the least-squares solution to a system of equations. The SVD is a common method to compute the pseudoinverse: given the SVD $X = U \Sigma V^*$, the pseudoinverse is $X^\dagger = V \Sigma^{-1} U^*$.

**Reduced-order model (ROM)** – A model of a high-dimensional system in terms of a low-dimensional state. Typically, an ROM balances accuracy with computational cost of the model.

**Regression** – A statistical model that represents an outcome variable in terms of indicator variables. Least-squares regression is a linear regression that finds the line of best fit to data; when generalized to higher dimensions and multilinear regression, this generalizes to PCR. Nonlinear regression, dynamic regression, and functional or semantic regression are used in system identification, model reduction, and machine learning.

**Restricted isometry property (RIP)** – The property that a matrix acts like a unitary matrix, or an isometry map, on sparse vectors. In other words, the distance between any two sparse vectors is preserved if these vectors are mapped through a matrix that satisfies the RIP.

**Singular value decomposition (SVD)** – Given a matrix $X \in \mathbb{C}^{n \times m}$, the SVD is given by $X = U \Sigma V^*$, where $U \in \mathbb{C}^{n \times n}$, $\Sigma \in \mathbb{C}^{n \times m}$, and $V \in \mathbb{C}^{m \times m}$. The matrices $U$ and $V$ are unitary, so that $UU^* = U^*U = I$ and $VV^* = V^*V = I$. The matrix $\Sigma$ has entries along the diagonal corresponding to the singular values ordered from largest to smallest. This produces a hierarchical matrix decomposition that splits a matrix into a sum of rank-one matrices given by the outer product of a column vector (left singular vector) with a row vector (conjugate transpose of right singular vector). These rank-one matrices are ordered by the singular value so that the first $r$ rank-one matrices form the best rank-$r$ matrix approximation of the original matrix in a least-squares sense.
**Snapshot** – A single high-dimensional measurement of a system at a particular time. A number of snapshots collected at a sequence of times may be arranged as column vectors in a data matrix.

**Sparsity** – A vector is sparse if most of its entries are zero or nearly zero. Sparsity refers to the observation that most data are sparse when represented as vectors in an appropriate transformed basis, such as a Fourier or POD basis.

**State-space** – The set of all possible system states. Often the state-space is a vector space, such as \( \mathbb{R}^n \), although it may also be a smooth manifold \( \mathcal{M} \).

**System identification** – The process by which a model is constructed for a system from measurement data, possibly after perturbing the system.

**Time-delay coordinates** – An augmented set of coordinates constructed by considering a measurement at the current time along with a number of times in the past at fixed intervals from the current time. Time-delay coordinates are often useful in reconstructing attractor dynamics for systems that do not have enough measurements, as in the Takens embedding theorem.

**Total least squares** – A least-squares regression algorithm that minimizes the error on both the inputs and the outputs. Geometrically, this corresponds to finding the line that minimizes the sum of squares of the total distance to all points, rather than the sum of squares of the vertical distance to all points.

**Uncertainty quantification** – The principled characterization and management of uncertainty in engineering systems. Uncertainty quantification often involves the application of powerful tools from probability and statistics to dynamical systems.

**Underdetermined system** – A system \( \mathbf{A} \mathbf{x} = \mathbf{b} \) where there are fewer equations than unknowns. Generally the system has infinitely many solutions \( \mathbf{x} \) unless \( \mathbf{b} \) is not in the column space of \( \mathbf{A} \).

**Unitary matrix** – A matrix whose complex conjugate transpose is also its inverse. All eigenvalues of a unitary matrix are on the complex unit circle, and the action of a unitary matrix may be thought of as a change of coordinates that preserves the Euclidean distance between any two vectors.
Bibliography


Index

action potential, 185
adjoint, 27, 28, 32, 40–43, 45, 101
algorithmic trading, 195, 196
Arnoldi algorithm, 10
background subtraction, 58, 70, 75, 89
balanced proper orthogonal decomposition, 32, 101, 109
coherent structures, 1, 25, 26, 126
composition, 7, 44
compressed sensing, 21, 133–136, 138–142
computer vision, 1, 55, 56
continuous wavelet transform (CWT), 74
control, 23, 32, 91–93, 95, 96, 98, 100, 101, 109, 110, 117, 184
correlation theory, 3, 134
delay coordinates, 9, 24, 105, 108, 110–112, 121, 189
diagnostics, 2
dynamical system, 1–4, 6–8, 18, 19, 21–25, 29, 30, 37, 39, 41, 43–49, 51, 52, 56, 78, 80, 87, 91, 97–102, 128, 130, 133, 138, 143, 144, 159, 160, 167, 174, 180, 187, 188, 196
EEG, 187
eigendecomposition, 4, 6–8, 23, 29, 79, 82, 94, 95, 102
eigenfunctions, 8, 39
eigensystem realization algorithm (ERA), 23, 24, 101, 105, 109–112, 139
eigenvalue, 8, 40, 42, 44
El Niño, 85–87
empirical orthogonal functions, 2, 22
epidemiology, 9, 100, 178
feature space, 160, 166, 167, 169, 170, 172
financial trading, 195–197, 202
flow around a cylinder, 25, 33, 35, 125, 127, 139, 148
fluid dynamics, 1, 9, 25, 27, 31, 32, 101, 133, 179
Gábor transform, 71, 72
Galerkin projection, 29, 30
Hankel matrices, 109–113, 120, 139
hidden Markov models, 105, 113–115, 117
Hilbert space, 6, 7, 39, 42, 43, 101
independent component analysis, 11
infectious disease, 91, 92, 100, 177–181, 184
Koopman eigenfunction, 44, 45, 50, 102, 160, 164, 166, 167
Koopman eigenvalues, 45, 46, 48, 164, 166
Koopman mode, 44–46, 102, 160, 167
Koopman spectra, 31, 39, 48, 100, 101
Koopman theory, 6, 7, 39, 41, 43–45, 47–49, 51, 163, 174
Krylov subspace, 10
manifold, 20, 43, 51, 52, 101, 159
manifold learning, 159
Moore–Penrose pseudoinverse, 5
mother wavelet, 73
multiresolution analysis, 72, 87
Navier–Stokes equation, 25, 27, 29, 30, 33
neuron, 185–187
neuroscience, 9, 19, 187
noise, 9, 67, 80, 105, 120, 121, 126–129, 136, 143–145, 187, 188
nonlinear Schrödinger equation, 161, 163, 172
Nyquist, 133, 134, 136, 137, 139, 140
particle image velocimetry (PIV), 27, 31, 133, 139, 140
Perron–Frobenius operator, 33, 45
polynomial kernel, 170, 171, 174
Koopman eigenfunctions, 8, 39
Koopman eigenvalues, 45, 46, 48, 164, 166
Koopman mode, 44–46, 102, 160, 167
Koopman spectra, 31, 39, 48, 100, 101
Koopman theory, 6, 7, 39, 41, 43–45, 47–49, 51, 163, 174
Krylov subspace, 10
manifold, 20, 43, 51, 52, 101, 159
manifold learning, 159
Moore–Penrose pseudoinverse, 5
mother wavelet, 73
multiresolution analysis, 72, 87
Navier–Stokes equation, 25, 27, 29, 30, 33
neuron, 185–187
neuroscience, 9, 19, 187
noise, 9, 67, 80, 105, 120, 121, 126–129, 136, 143–145, 187, 188
nonlinear Schrödinger equation, 161, 163, 172
Nyquist, 133, 134, 136, 137, 139, 140
particle image velocimetry (PIV), 27, 31, 133, 139, 140
Perron–Frobenius operator, 33, 45
polynomial kernel, 170, 171, 174
power spectrum, 119–124, 137, 190, 192
principal component analysis, 2, 11, 23, 29, 135
principal component pursuit, 56, 57
proper orthogonal decomposition, 2, 6, 25–32, 119, 135
pseudoinverse, 5, 7, 8, 11, 125, 167–169
radial basis functions, 160, 171, 174
reduced-order modeling, 5, 24, 32, 101
regression, 2, 5, 6, 8, 21, 23, 24, 128, 131
restricted isometry property, 135, 140, 142, 143
robust principal component analysis, 55
sampling rate, 78, 133, 136, 139
Shannon–Nyquist, 134, 136, 137, 139
shift-stacking, 105, 108
sigmoid kernel, 171
singular value decomposition, 2, 7, 14–17, 20, 26, 28, 29, 35, 55, 77–80, 87–89, 93, 95, 96, 98, 119, 121, 126, 127, 131, 135, 136, 142, 143, 148, 166, 167, 169, 170, 197, 204
SIR models, 179
sparsity, 24, 66, 124, 133, 134, 136–140, 143, 145, 149
state prediction, 1, 2, 4, 8, 19–21, 32, 48, 166
stock market, 195–197, 199, 200
support vector machine, 160, 170
system identification, 5, 24, 32, 93, 100, 101
thresholding, 7, 63, 69, 126
time-frequency analysis, 72, 74
truncation, 7, 15, 17, 24, 32, 48, 66, 80, 93–96, 109, 110, 119, 121, 126, 127
uncertainty quantification, 21, 27, 33, 45
unitary, 135, 139, 140, 142, 143
vaccination, 92, 100
video surveillance, 55, 56, 63
volatility, 195
wavelets, 73–75, 78, 135
Data-driven dynamical systems is a burgeoning field—it connects how measurements of nonlinear dynamical systems and/or complex systems can be used with well-established methods in dynamical systems theory. This is a critically important new direction because the governing equations of many problems under consideration by practitioners in various scientific fields are not typically known. Thus, using data alone to help derive, in an optimal sense, the best dynamical system representation of a given application allows for important new insights. The recently developed dynamic mode decomposition (DMD) is an innovative tool for integrating data with dynamical systems theory. The DMD has deep connections with traditional dynamical systems theory and many recent innovations in compressed sensing and machine learning.

Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems, the first book to address the DMD algorithm,

• presents a pedagogical and comprehensive approach to all aspects of DMD currently developed or under development;
• blends theoretical development, example codes, and applications to showcase the theory and its many innovations and uses;
• highlights the numerous innovations around the DMD algorithm and demonstrates its efficacy using example problems from engineering and the physical and biological sciences; and
• provides extensive MATLAB code, data for intuitive examples of key methods, and graphical presentations.

The core audience for this book is engineers and applied mathematicians working in the physical and biological sciences. It can be used in courses that integrate data analysis with dynamical systems.

J. Nathan Kutz is the Robert Bolles and Yasuko Endo Professor of Applied Mathematics, Adjunct Professor of Physics and Electrical Engineering, and Senior Data Science Fellow with the eScience Institute at the University of Washington, Seattle.

Steven L. Brunton is an Assistant Professor of Mechanical Engineering, Adjunct Assistant Professor of Applied Mathematics, and a Data Science Fellow with the eScience Institute at the University of Washington, Seattle.

Bingni W. Brunton is the Washington Research Foundation Innovation Assistant Professor of Biology and a Data Science Fellow with the eScience Institute at the University of Washington, Seattle.

Joshua L. Proctor is an Associate Principal Investigator with the Institute for Disease Modeling as well as Affiliate Assistant Professor of Applied Mathematics and Mechanical Engineering at the University of Washington, Seattle.

For more information about SIAM books, journals, conferences, memberships, or activities, contact:

Society for Industrial and Applied Mathematics
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
+1-215-382-9800 • Fax +1-215-386-7999
siam@siam.org • www.siam.org