With state feedback, we have

$$x_{t+1} = Ax_t + Bu_t = Ax_t + B(Kx_t) = (A + BK)x_t, \quad t = 1, 2, \ldots.$$

This recursion is called the *closed-loop system*. The matrix $A + BK$ is called the *closed-loop dynamics matrix*. (In this context, the recursion $x_{t+1} = Ax_t$ is called the *open-loop system*. It gives the dynamics when $u_t = 0$.)

10.3 Matrix power

It makes sense to multiply a square matrix A by itself to form AA. We refer to this matrix as A^2. Similarly, if k is a positive integer, then k copies of A multiplied together is denoted A^k. If k and l are positive integers, and A is square, then $A^kA^l = A^{k+l}$ and $(A^k)^l = A^{kl}$. By convention we take $A^0 = I$, which makes the formulas above hold for all nonnegative integer values of k and l.

We should mention one ambiguity in matrix power notation that occasionally arises. When A is a square matrix and T is a nonnegative integer, A^T can mean either the transpose of the matrix A or its Tth power. Usually which is meant is clear from the context, or the author explicitly states which meaning is intended. To avoid this ambiguity, some authors use a different symbol for the transpose, such as A^{T} (with the superscript in roman font) or A', or avoid referring to the Tth power of a matrix. When A is not square there is no ambiguity, since A^T can only be the transpose in this case.

Other matrix powers. Matrix powers A^k with k a negative integer will be discussed in §11.2. Non-integer powers, such as $A^{1/2}$ (the matrix squareroot), need not make sense, or can be ambiguous, unless certain conditions on A hold. This is an advanced topic in linear algebra that we will not pursue in this book.

Paths in a directed graph. Suppose A is the $n \times n$ adjacency matrix of a directed graph with n vertices:

$$A_{ij} = \begin{cases} 1 & \text{there is a edge from vertex } j \text{ to vertex } i \\ 0 & \text{otherwise} \end{cases}$$

(see page 112). A *path* of length ℓ is a sequence of $\ell + 1$ vertices, with an edge from the first to the second vertex, an edge from the second to third vertex, and so on. We say the path goes from the first vertex to the last one. An edge can be considered a path of length one. By convention, every vertex has a path of length zero (from the vertex to itself).

The elements of the matrix powers A^ℓ have a simple meaning in terms of paths in the graph. First examine the expression for the i, j element of the square of A:

$$(A^2)_{ij} = \sum_{k=1}^{n} A_{ik}A_{kj}.$$

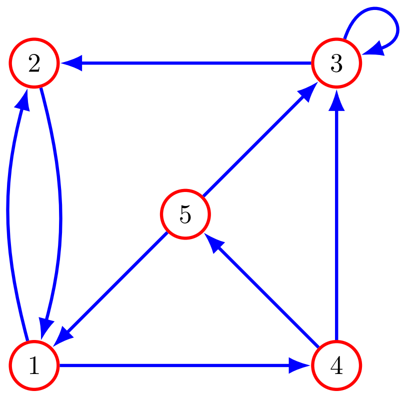

Figure 10.1 Directed graph.

Each term in the sum is 0 or 1, and equal to one only if there is an edge from vertex j to vertex k and an edge from vertex k to vertex i, *i.e.*, a path of length exactly two from vertex j to vertex i via vertex k. By summing over all k, we obtain the total number of paths of length two from j to i.

The adjacency matrix A for the graph in figure 10.1, for example, and its square are given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \qquad A^2 = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We can verify there is exactly one path of length two from vertex 1 to itself, *i.e.*, the path $(1,2,1)$), and one path of length two from vertex 3 to vertex 1, *i.e.*, the path $(3,2,1)$. There are two paths of length two from vertex 4 to vertex 3, $(4,3,3)$ and $(4,5,3)$, so $(A^2)_{34} = 2$.

The property extends to higher powers of A. If ℓ is a positive integer, then the i,j element of A^ℓ is the number of paths of length ℓ from vertex j to vertex i. This can be proved by induction on ℓ. We have already shown the result for $\ell = 2$. Assume that it is true that the elements of A^ℓ give the paths of length ℓ between the different vertices. Consider the expression for the i,j element of $A^{\ell+1}$:

$$(A^{\ell+1})_{ij} = \sum_{k=1}^{n} A_{ik}(A^\ell)_{kj}.$$

The kth term in the sum is equal to the number of paths of length ℓ from j to k if there is an edge from k to i, and is equal to zero otherwise. Therefore it is equal to the number of paths of length $\ell+1$ from j to i that end with the edge (k,i), *i.e.*, of the form $(j,\ldots,k,i)$. By summing over all k we obtain the total number of paths of length $\ell+1$ from vertex j to i.

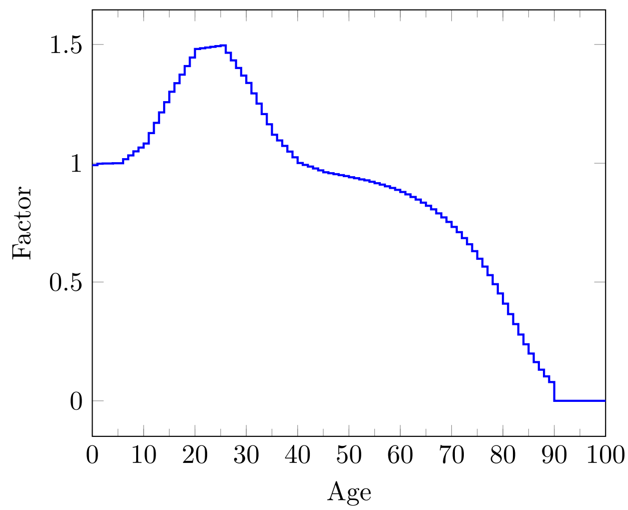

Figure 10.2 Contribution factor per age in 2010 to the total population in 2020. The value for age $i-1$ is the ith component of the row vector $\mathbf{1}^T A^{10}$.

This can be verified in the example. The third power of A is

$$A^3 = \begin{bmatrix} 1 & 1 & 1 & 1 & 2 \\ 2 & 0 & 2 & 3 & 1 \\ 2 & 1 & 1 & 2 & 2 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

The $(A^3)_{24} = 3$ paths of length three from vertex 4 to vertex 2 are $(4,3,3,2)$, $(4,5,3,2)$, $(4,5,1,2)$.

Linear dynamical system. Consider a time-invariant linear dynamical system, described by $x_{t+1} = Ax_t$. We have $x_{t+2} = Ax_{t+1} = A(Ax_t) = A^2x_t$. Continuing this argument, we have

$$x_{t+\ell} = A^\ell x_t,$$

for $\ell = 1, 2, \ldots$. In a linear dynamical system, we can interpret A^ℓ as the matrix that propagates the state forward ℓ time steps.

For example, in a population dynamics model, A^ℓ is the matrix that maps the current population distribution into the population distribution ℓ periods in the future, taking into account births, deaths, and the births and deaths of children, and so on. The total population ℓ periods in the future is given by $\mathbf{1}^T(A^\ell x_t)$, which we can write as $(\mathbf{1}^T A^\ell)x_t$. The row vector $\mathbf{1}^T A^\ell$ has an interesting interpretation: Its ith entry is the contribution to the total population in ℓ periods due to each person with current age $i-1$. It is plotted in figure 10.2 for the US data given in §9.2.

Matrix powers also come up in the analysis of a time-invariant linear dynamical system with an input. We have

$$x_{t+2} = Ax_{t+1} + Bu_{t+1} = A(Ax_t + Bu_t) = A^2 x_t + ABu_t + Bu_{t+1}.$$

Iterating this over ℓ periods we obtain

$$x_{t+\ell} = A^\ell x_t + A^{\ell-1}Bu_t + A^{\ell-2}Bu_{t+1} + \cdots + Bu_{t+\ell-1}. \qquad (10.2)$$

(The first term agrees with the formula for $x_{t+\ell}$ with no input.) The other terms are readily interpreted. The term $A^j Bu_{t+\ell-j}$ is the contribution to the state $x_{t+\ell}$ due to the input at time $t+\ell-j$.

10.4 QR factorization

Matrices with orthonormal columns. As an application of Gram matrices, we can express the condition that the n-vectors $a_1, \ldots, a_k$ are orthonormal in a simple way using matrix notation:

$$A^T A = I,$$

where A is the $n \times k$ matrix with columns $a_1, \ldots, a_k$. There is no standard term for a matrix whose columns are orthonormal: We refer to a matrix whose columns are orthonormal as 'a matrix whose columns are orthonormal'. But a *square* matrix that satisfies $A^T A = I$ is called *orthogonal*; its columns are an orthonormal basis. Orthogonal matrices have many uses, and arise in many applications.

We have already encountered some orthogonal matrices, including identity matrices, 2-D reflections and rotations (page 129), and permutation matrices (page 132).

Norm, inner product, and angle properties. Suppose the columns of the $m \times n$ matrix A are orthonormal, and x and y are any n-vectors. We let $f : \mathbf{R}^n \to \mathbf{R}^m$ be the function that maps z to Az. Then we have the following:

- $\|Ax\| = \|x\|$. That is, f is *norm preserving*.
- $(Ax)^T(Ay) = x^T y$. f preserves the inner product between vectors.
- $\angle(Ax, Ay) = \angle(x, y)$. f also preserves angles between vectors.

Note that in each of the three equations above, the vectors appearing in the left- and right-hand sides have different dimensions, m on the left and n on the right.

We can verify these properties using simple matrix properties. We start with the second statement, that multiplication by A preserves the inner product. We have

$$\begin{aligned}(Ax)^T(Ay) &= (x^T A^T)(Ay)\\ &= x^T(A^T A)y\\ &= x^T I y\\ &= x^T y.\end{aligned}$$